

FIGURE 168

GTGGAGGCCGCGACGATGGCGGGCCGACGGAGGCCGAGACGGGGTTGGCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCACGCCGCTGGGTGTTCTGCTCGCATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTGCACCTGTGGCTGACGTCA
TGCTGAGGACTTGGTCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACCTTGCCGGAGTGTGCTACGCATGGTGCCTGCATGGT
TGTGGGACCCAAAACCCATTGCCTTCATGGGTGGCCAGAGCCTCTGTGCCCTGCC
AGAGCCTGGTCATCTTCTCCAGCCAAGCTGGCTGCCTGTGGTCCAGAGCACCAGCGA
GCCACGGCCAACATGCTGCCACCATGTCGAACCCCTGGCGTCCTGTGGCCAATGTGCT
GTCCCCCTGTGCTGGTCAAGAACGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCTGCCTGTCCACCATCTGCCGTGTCGGAGAGTGTGCCCTGGGGGG
CCCTCTGCCGGGGCTGCCAGCTCACCTCAGAGAACGTTCTGGATGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGCTTGGGGAAATGATCGGATCTCTG
CCAGCTCTCAGCCCTCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGTTTCC
GGCCTCTGTGGCGCTCTTCATCACGTTGGATCCTGGGGACTGGCTCTCGGCCCTA
TGTGGACCGGACCAAGCAACTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTGG
CCTGCGTGCCCTTGCCTGGTGTCCCAGCTGCAGGGACAGACCCCTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGCTGTTGGCTCTCGGTGGGCCATGGAGTTGGCGGTGCA
GTGTTCTTCCCCGTGGGGAGGGGGCTGCCACAGGCATGATCTTGTGCTGGGCAGGCCG
AGGGAATACTCATGCTGGCAATGACGGCACTGACTGTGCGACGCTGGAGCCGTCTTG
TCCACCTGCCAGCAGGGGAGGATCCACTGACTGGACAGTGTCTGCTGCTGATGGCCGG
CCTGTCACCTTCTCAGCTGCATCCTGGCGTCTTCCACACCCCATACGGCGCCTGC
AGGCCGAGTCTGGGAGCCCCCTCCACCGTAACGCCGTGGCGGCCAGACTCAGGCCG
GGTGTGGACCGAGGGGAGCAGGAAGGGCTGGGTCTGGGCCAGCACGGCACTCCGGA
GTGCACGGCGAGGGGGCCTCGCTAGAGGACCCAGAGGGCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCCCTCCGCCCG
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTATTGACCCGGCTGGGTCTACTCCTCCTT
CTCCTCCCCGTGGGTGATCACGTAGCTGAGCGCCTGTAGTCCAGGTTGCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCGGGCGGGGGCAAAGGGCTCCTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

170/237

FIGURE 169

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCKMVVTQ
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRTANMLATMSNPLGVLVANVLSPLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNK
YVILAVCLGGMIGISASFASALLEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLGQAEGILIMLAMTALTVRSEPSLSTCQQGEDPLDWTVSLLL MAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAA TDAPS RPGRLAGR VQASRFIDPAGSHSSFSSPWVIT
```

Important features:

Potential Transmembrane domains:

amino acids 30-50, 61-79, 98-112, 126-146, 169-182, 201-215, 248-
268, 280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

171/237

FIGURE 170A

GTCCCACATCCTGCTCACTGGGTCAAGGTCCCTCTTAGACCAGCTTGTCCATTTGCT
 GAAGTGGACCAACTAGTCCCCAGTAGGGGGTCTCCCTGGCAATTCTGATCGCGTTGG
 ACATCTCAGATCGCTTCCAATGAAGATGGCCTGCCCTGGCTCTGGGGCTGTTCATATACTA
 TCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGAAAGGAGCACGGGGCTGATCAAGC
 CATCCAGGAAACACTGGAGGACTTGTCCAGCCTGAAAGAACTCTAGTGGTTCTGAATCTA
 GCCCACTTGGCGGTAAAGCATGATGCAACTCTGCAACTTCTGCTGGGGCTTTGGGCCAGG
 TGGCTACTTATTCTTTAGGGATTGTCAAGGAGGTGACCACTCTCACGGTCAAATACCAAG
 TGTCAGAGGAAGTGCATCTGGTACAGTGATCGGGAAAGCTGTCAGGAACTGGGCCGGAG
 GAGAGGCCGGAGGCAAGCTGGGGCCGCTTCCAGGTGTCAGCTGCCAGGCCCTCCCCAT
 TCAGGTGGACTCTGAGGAAGGCTGCTCACAGGCAGGCCGGCTGGATCGAGAGCAGCTGT
 GCCGACAGTGGGATCCCTGCCCTGGTTCCCTTGATGTGCTGCCACAGGGATTGGCTCTG
 ATCCATGTGGAGATCCAAGTGTGGACATCAATGACCAACAGCCACGGTTCCCAAAGGCAG
 GCAGGAGCTGGAAATCTTGAGAGCGCCTCTCGCAACCCGGATCCCCCTGGACAGAGCTC
 TTGACCCAGACACAGGCCAACACCCCTGCACACACTCTGTCTCCAGTGAGCAGCTTT
 GCCTTGGATGTCATTGGGGCCCTGATGAGACCAAACATGCAGAACTCATAGTGGTGAAGGA
 GCTGGACAGGGAAATCCATTGATCTGGTGTAACTGCCTATGACAATGGGAACCC
 CCCCCAAGTCAGGTACAGCTTGTCAAGGTCAACGTCTGGACTCCAATGACAATAGCCCT
 GCGTTGCTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCT
 CATAAAACTGACCGCCACAGACCTGACCAAGGCCCAATGGGAGGTGGAGTTCTCCTCA
 GTAAGCACATGCCCTCAGAGGTGTCAGCCTCAGTATTGATGCCAAGACAGGCCAGGTC
 ATTCTGCGTCGACCTCTAGACTATGAAAAGAACCCCTGCCTACGAGGGATGTTCAAGGCAAG
 GGACCTGGGTCCAATCCTATCCAGCCCATTGCAAAGTTCTCATCAAGGTTCTGGATGTC
 ATGACAACATCCAAGCATCCACGTACATGGGCCCTCAGCCATCAGTGGTGTCAAGAGCT
 CTTCCCAAGGACAGTTATTGCTCTGTGATGGCAGATGACTTGGATTCAAGGACACAATGG
 TTTGGTCCACTGCTGGTGAGCCAAGAGCTGGCCACTTCAGGCTGAAAGAAACTATGGCA
 ACACATACATGTTGCTAACCAATGCCACACTGGACAGAGCAGTGGCCCAAATACCC
 ACTCTGTTAGCCAAGACCAAGGACTCCAGCCCTATGCCAAGAAACAGCTCAGCATTCA
 GATCAGTGACATCAACGACAATGCCACCTGTGTTGAGAAAAGCAGGTATGAAGTCTCCACGC
 GGGAAAACAACCTACCCCTCTTCACCTCATTACCATCAAGGCTCATGATGCAGACTGGGC
 ATTAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTGCTCACTTAGCTATTGA
 CTCCAACACAGGAGAGGTCACTGCTCAGAGGTCACTGAACATATGAAGAGATGCCGGCTTG
 AGTTCCAGGTGATCGCAGAGGACAGCAGGGCAACCCATGCTTCAGTGTCTCTGTGTTGG
 GTCAGCCTCTGGATGCCAATGATAATGCCCAAGAGGTGGCCAGCCTGTGCTCAGCGATGG
 AAAAGCCAGCCTCTCCGTGTTGTAATGCCCTCACAGGCCACCTGCTGGTCCCCTGAGA
 CTCCCAATGGCTGGGCCAGCGGGCACTGACACACCTCCACTGCCACTCACAGCTCCGG
 CCATTCCCTTGACAACCATGTTGCAAGAGATGCAGACTGGGGCAAATGGAGAGCC
 CTACAGCATCCGAATGAAATGAAGCCCACCTCTCATCCTCAACCCCTCATACGGGCAGC
 TGTTGTCATGTCACCAATGCCAGCAGCCTCATTGGAGTGAGTGGAGCTGGAGATAGTA
 GTAGAGGACCAAGGAAAGCCCCCTTACAGACCCGAGGCCCTGAGGGTCATGTTGTCAC
 CAGTGTGGACCACTGAGGGACTCAGCCGCAAGCCTGGGCCCTGAGCATGTCATGCTGA
 CGGTGATCTGCCCTGGCTGACTGTTGGCATCTCGGGTTGATCTGGCTTTGTCATGTC
 ATCTGCCGGACAGAAAAGAAGGACAACAGGGCTACAACACTGTCGGAGGCCAGTCCACCTA
 CCGCCAGCAGCCCAGAGGCCCAAGGGCCAGAAACACATTGAGACAGACATCCACCTCGTGCCTG
 TGCTCAGGGGTCAAGGAGGTGAGCCTTGTAAGTGCAGACTGGGCCAGTCCACAAAGATGTGGACAAG
 GAGGCAGTGTGAAAGCAGGCTGGACCCCTGCCCTGCAAGGCCCTTCCACCTCACCCGAC
 CCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGGAGCAGGCCAGAGGCCAGAGGTG
 TGCAAGACACGGTCAACCTCTTCAACCCTCCAGGCAAGGAGGAATGCCCTCCGGAGAAG
 CTGAAACCTTCCCGAGGCCAGCCTGCCACAGGCCAGCTCCAGGGCTCTGAAGGGTGC
 AGGCAGCCCCACAGGGAGGCTGGCTGGAGACCAGGGCAGTGAGGAAGGCCACAGAGGCCAC
 CAGCCTCTGCAACCTGAGACGGCAGCGACATCTCAATGGCAAAGTGTCCCTGAGAAA
 GAATCAGGGCCCCGTCAAGATCCTGCGGAGCCTGGTCCGGCTGTGTCAGGAGCTG
 GCGGAACCCCGTGGAGGAGCTCACTGTGGATTCTCCTCTGTCAGCAAATCTCCAGCTGC
 TGTCCTGCTGCACTAGGGCAATTCCAGCCAAACCAACCCACAGGAGGAAATAAGTACTTG
 GCCAAGCCAGGAGGAGCAGCAGGAGTGCAATCCAGACACAGATGCCCAAGTGCAAGGGCTGG

172/237

FIGURE 170B

AGGCCAGACAGACCCAGAACAGGAGGAAGGGCTTGGATCCTGAAGAGGACCTCTGTGA
AGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCCCAGCACAGGTCTGGCCCTGGAC
CGGCTGAGCGCCCCCTGACCCGGCTGGATGGCGAGACTCTCTTGCCCCCTCACCAACTA
CCGTGACAATGTGATCTCCCCGGATGCTGCAGCCACGGAGGAGGCCAGGACCTTCAGACGT
TCGGCAAGGCAGAGGCACCAGAGCTGAGCCAACAGGCACGAGGCTGGCCAGCACCTTGTC
TCGGAGATGAGCTCACTGCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCGTGGAGGC
CGCCTCCGAGGCCTGCAGGGCTCTCGGTCTGCGGGAGGACCCCTCAGTTAGACTTGGCCA
CCAGTGCAGCCTCAGGCATGAAAGTGAAGGGGACCCAGGTGGAAAGACGGGACTGAGGGC
AAGAGCAGAGGCAGCAGCAGCAGCAGGTGCCTGTGAACATACTCAGACGCCTCTGGAT
CCAAGAACCAAGGGCCTGAGGATCTGTGGACAAGAGCTGGTTCTAAAATCTTGTAACTCAC
TAGCTAGCGGCCTGAGAACTTCTGGACTGATGCTACCCCCACAGAGGAGGCAAGAG
CCCCAGGACTAACAGCTGACTGACCAAAGCAGCCCTGTAAGCAGCTTGAGTCTTGGA
GGACAGGGACGGTTCTGGCTGAGATAAGTGTCTGGAAAACATATGTGGAGCACAAAG
GGTCAGTCCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGGAAAGGGTGGCCTTCTT
GGGTAGCAGGAGTCAGGGGCTGTACCTGGGGTGCAGGAAATGCTCTGACCTATCAA
TAAAGGAAAAGCAGTAAAAAAAAAAAAAA

173/237

FIGURE 171

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLLQLLLGLLPGGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTIVGKLSQELGREERRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHVEIQ
VLDINDHQPRFPKGQELEISESASLRTRIPLDRA LD PDTGPNTLHTYTLSPSEH FAL DVIV
GPDET KHAELIVVKEL DREIHSFFDLVLTAYDNGNPPSGTSLVKVNVLDSNDNSPAFAESS
LA LEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEV DVQARDLGP NPIPAHCKVLI KVLDVNDNIPSIHVTWASQPSL VSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFR LKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHADLGIN GKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPM LASSVSVWSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPPLATHSSRPFLLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSDHRLDSARKPGALSMSMLTVICLAVLLGIFGLL ALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCLQAFHLLPTLYRTL RNQGNQGAPAESREV LQDTVNL FNHPRQRNASREN NLPEP
QPATGQPRS RPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVEELTVDSPPVQQI SQLSLLHQGQFQPKPNHRGNKYLAKPGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEE DL SVKQLLEEELSSLLDPSTGLALDRILSAPD
PAWMARLSSLPLTNYRDNVISPDAAAATEEPRTFQTFGKAEAPELSPTGTR LASTFVSEMSSL
LEM LLEQRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGD PGK TGTEGKSRGSS
SSSRCL
```

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

174/237

FIGURE 172

CGGACGCGTGGCGGACGCGTGGGGAGAGCCGCAGTCCCCTGCAGCACCTGGAGAAGG
CAGACCGTGTGAGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTGGGAGTGGAAAGTGGAG
GCAGGAGCCCTCCTACACTTCGCCATGAGTTCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTGGATTTGGCTTCTCATGCCAACATTGTTAAAGACTAT
GAGATACGTCACTATGTTACAGGTGATCTCTCCGTGACGTTGCATTTCTGCACCCT
GTTGAGCTCATCATTTGAAATCTTAGGAGTATTGAATAGCAGCTCCGTTATTTCACT
GGAAAATGAACCTGTGTGAATTCTGCTGATCCTGGTTTATGGTGCCTTTACATTGGC
TATTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTCTGTCTCTT
ATGGCTGACCTTATGTATTTCTGGAAACTAGGAGATCCCTTCCCATTCTCAGCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGTTGGTGTATTGGAGTGACTCTC
ATGGCTCTTCTTCTGGATTTGGTGTCAACTGCCATACACTTACATGTCTTACTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACATGTTCCAGAAGGGGAA
GTGCATAACAAACCATCAGGTTCTGGGAATGATAAAAGTGTACCTCAGCATCAGG
AAAGTAAAAATCTTACTCTTATTCAACAGGAAGTGGATGCTTGGAAAGAATTAAGCAGGCAGC
TTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAACACTCCAAAACCTTC
AAGGGAAATATTTAATTCTGGTTACTTTCTCTATTACTGTGTTGGAAAATTT
CATGGCTACCATCAATATTGTTGATCGAGTTGGAAAACGGATCCTGTCACAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGTGAAGTTGGTCCAACACATT
TCCTTCATTCTTGTGGAAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTATGCCATCTCAGCAGTAAGTCCTCAATGTCATTGCTGCTATTAGCAC
AGATAATGGCATGTACTTGTCTCCTGTGCTGATCCGAATGAGTATGCCTTAGAA
TACCGCACCATAATCACTGAAGTCCTGGAGAACTGCAGTTCAACTCTATCACCGTTGGTT
TGATGTGATCTTCTGGTCAGCGCTCTCTAGCATACTCTTCTATTGGCTCACAAAC
AGGCACCAGAGAACAAATGGCACCTTGAACCTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTCAAAATTAGATATAAGAGGGGGAAAATGGAACCAGGGCTGACATTATAAAC
AAACAAAATGCTATGGTAGCATTTCACCTTCATAGCATACTCCTCCCCGTAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAAGAACACTCA
GCAGAGAGCATCCGTGGATATGAGGCTGGTAGAGGGGGAGAGGAGCCAAGAAACTAA
AGGTAAAAATACACTGGAACCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTCCGTTTAAGGTTACATGGAAAAGGTTAGCTTGCCTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCG
ACCTGCAGAAGCTGGCCGCCATGCCAACCTGTTATTGCAGCTTATAATG

175/237

FIGURE 173

MSFLIDSSIMITSQILFFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPPILSPKPHGILSIEQLISRVGVIGVTLALLSGFGAVNCPTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWMKSVTTSASGSENLTLIQ
QEVDALEELSRLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLLAQIMGMYFVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

176/237

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTCCTTACACTGCCATGAGTTCCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTGATTTGGCTTCNGCGCCAATGTT
TAAAGACTATGAGATACTGTCAGTATGTTGATCNGGTGATCTCTCCGTGACGTTGCCATTT
CTTGCACCATGTTGAGCTCATCATCTTGAAATCTTNGGAGTATTGAATAGCAGCTCCGT
TATTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTT
CCTGTCTCTTATGGCTGACCTTATGTATTCCAG

177/237

FIGURE 175

GTGTTGCCCTGGGGAGGGGAAGGGGAGCCNGCCCTTCCTAAAATTGGCCAAGGGTTTC
TTTNTTGAATTCCGGGTTNNGNATACCTTCCCAGAAAATTTTTGGATTGGGTAGNTT
TTTTCATGCGCCAATTGTTAAAGACTATGAGATACTGTCAGTATGTTGACAGGTGATNTT
NTCCGTGACGTTGCATTTCTGCACCAGTTGAGCTCATCATNTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCGTTATTTCACTGGAAAATGAACCTGTGTAATTCTGCTGATC
CTGGTTTCATGGTGCCTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTCTGTCTNTTATGGCTGACCTTATGTATTNTNTGGAAAN
TAGGAGATCCCTTCCCATTCTC

178/237

FIGURE 176A

CTCGCGCAGGGATCGTCCCATGGCCGGGCTCGGAGCCGCACCCCTGGGGGGCCTCCGGGA
 TTTGCTACCTTTGGCTCCCTGCTCGTAACGTCTCTCACGGCTGTCGCCCTCAAT
 CTGGACGTGATGGGTGCCTTGCAGCAAGGAGGGCGAGCCAGGCAGCCTCTGGCTTCTGT
 GGCCTGCACCGGCAGTGCAGCCCCGACCCCAGAGCTGGCTGCTGGTGGGTGCTCCCCAGG
 CCTGGCTCTTCCGGCAGCAGCGAATCGCACTGGAGGCCTTCGCTTGCCCGTTGAGC
 CTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGAAAAGGAAAG
 CAAGGAGAACCAAGTGGTGGGAGTCAGTGTGAGCCAGGGCTGGGGCAAGATTGTTA
 CCTGTGCACACCGATATGAGGAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATG
 ATTGGTCGCTGCTTGTGCTCAGCCAGGACCTGGCATCCGGATGAGTTGGATGGTGGGA
 ATGGAAGTTCTGTGAGGGACGCCCAAGGCCATGAACAATTGGGTTCTGCCAGCAGGGCA
 CAGCTGCCGCCTCTCCCTGATAGCCACTACCTCTTGGGCCCCAGGAACCTATAAT
 TGGAAAGGGCACGCCAGGGTGGAGCTGTGACAGGGCTCAGCGGACCTGGCACACCTGGA
 CGACGGTCCCTACGAGGGGGGGAGAGAAGGAGCAGGACCCCCGCCTCATCCGGTCCCTG
 CCAACAGCTACTTGGCTCTCATTGACTGGGAAAGGCTGGTGTGGTATCTGCGCACAGGA
 AGCTTTGTGGCTGGAGCCCCCGCCAACCACAAGGGTGTGGTGTGGTATCTGCGCACAGGA
 CAGCGCCAGTCGCCCTGGTGCAGGTTATGCTGTGGGAGCGCCTGACCTCCGGCTTTG
 GCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGCCAGACCTGATAGTGGTGCC
 CCCTACTCTTGAGGCCAAGAAGAGCTGGGGGTGCTGTGTATGTGACTTGAACCAGGG
 GGGTCACTGGCTGGATCTCCCTCTCCGGCTCTGGCTCCCTGACTCCATGTTGGGA
 TCAGCCTGGCTGCTGGGACCTCAACCAAGATGGCTTCCAGATATTGCACTGGGTGCC
 CCCTTGATGGTGTGGAAAGTCTCATCTACCATGGGAGCAGCCTGGGGTTGTGCCAA
 ACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGATCAAGAGCTTCCGGCTACTCCCTGTCAG
 GCAGCTGGATATGGATGGGAACCAATACCCCTGACCTGCTGGTGGGCTCCCTGGCTGACACC
 GCAGTGCCTTCAGGGCCAGACCCATCTCCATGCTCCCATGAGGTCTCTATTGCTCCACG
 AAGCATCGACCTGGAGCAGCCAACTGTGCTGGCGGCCACTGGTCTGTGGACCTAAGGG
 TCTGTTCAGCTACATTGCACTGCCCCAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTG
 TTAGATGCGGACACAGACGGAGGCTCCGGGCCAGGTTCCCGTGTGACGTTCTGAGCCG
 TAACCTGGAAGAACCAAGCACCGGCTCGGGCACCGTGTGGCTGAAGCACCAAGCTCGGGCATT
 GTAGTGCACCTGTCCTACAGTCTCCAGACCCCTGGCTCCGGGACAGGCTCTGGCCAGGG
 GCTGCCTCCAGTGGCCCCATCTCAATGCCACCAGCCAGCAGCAGGGCAGAGATCC
 ACTTCCTGAAGCAAGGCTGTGGTGAAGACAAGATCTGCCAGAGCAATCTGCAGCTGGTCCAC
 GCCCGCTCTGTACCCGGGTAGCGACACGGAAATTCAAACCTCTGCCCATGGATGTGGATGG
 AACAAACAGCCCTGTTGCACTGAGTGGCAGCCAGTCATTGGCTGGAGCTGATGGTCACCA
 ACCTGCCATGGACCCAGCCCAGGGCTGATGGGATGATGCCCATGAAGGCCAGC
 CTGGTCATGCTTCTGACTCACTGCACTACTCAGGGTCCGGGCCCTGGACCCCTGCGGAGAA
 GCCACTCTGCCTGTCCAATGAGAATGCCCTCCATGTTGAGTGTGAGCTGGGAACCCATGA
 AGAGAGGTGCCAGGTACCTCTACCTCATCCTAGCACCTCCGGGATCAGCATTGAGACC
 ACGGAACCTGGAGGTAGAGCTGCTGGCCAGCATCAGTGGAGCAGGAGCTGCATCCAGTCTC
 TGCACGAGCCCGTGTCTCATGAGCTGCCACTGTCCATTGCAAGGAATGCCATTCCCCAGC
 AACTCTCTCTGGTGTGGTAGGGGGAGAGAGGCCATGCACTGAGCTGAGCGGGATGTGGG
 AGCAAGGTCAAGTATGAGGTACGGTTCCAACCAAGGCCAGTCGCTCAGAACCCCTGGGCTC
 TGCCTTCTCAACATCATGTCCTCATGAGATTGCCAATGGGAAGTGGTGTGTACCCAA
 TGCAAGGTTGAGCTGGAGGGCGGGCAGGGGCTGGCAGAAAGGGCTTGCTCTCCAGGCC
 AACATCCCTCACCTGGATGTGGACAGTAGGGATAGGGAGCGGGAGCTGGAGCCACCTGA
 GCAGCAGGGAGCCTGGTAGCGGAGGCCAGCATGTCCTGGTGGCCAGTGTCTCTGCTG
 AGAAGAAGAAAACATCACCCCTGGACTGCCCGGGGAGGGCCAACTGTGTGGTGTACGC
 TGCCCCACTCTACAGCTTGACCGCGCGCTGTGCTGCATGTCATGGGGCGTCTCTGGAAACAG
 CACCTTCTGGAGGGAGTACTCAGCTGTGAAGTCCCTGGAAAGTGTGTCGGGGCAACATCA
 CAGTGAAGTCCTCCATAAGAACCTGATGCTCCAGATGCCCTCACAGTGTCCCAGTGTG
 GTATACTGGACCCATGGCTGTGGTAGCACTGCTGGTGTGGCAAGAGGAGTGCCTGGGTCATCCTCTGGC
 TGACTGGCTGGCTGTGGTAGCACTGCTGGTGTGGCAAGGAGTGGGATTCT
 TCAAACGGGCGAAGCACCCGAGGCCACCGTGCCTGGAAAGATGGGATTCT
 GAAGACCGACAGCAGTTCAAGGAGGAGAAGACGGGACCATCCTGAGGAACAACTGGGGCAG

179/237

FIGURE 176B

CCCCCGGGCGGGAGGGCCCGGATGCACACCCCATTCTGGCTGCTGACGGGCATCCGAGCTGG
GCCCGATGGCATCCAGGGCAGGCACCGCCTAGGTTCCCATGTCCCAGCCTGGCCTGTGG
CTGCCCTCCATCCCTCCCCAGAGATGGCTCCTGGATGAAGAGGGTAGAGTGGCTGCTG
GTGTCGCATCAAGATTTGGCAGGATCGGCTTCAGGGCACAGACCTCTCCCACCCACAA
GAACTCCTCCCACCCAACTTCCCTTAGAGTGCTGAGATGAGAGTGGTAAATCAGGGAC
AGGGCCATGGGTAGGGTGAGAAGGGCAGGGTGTCTGATGCAAAGGTGGGAGAAGGGAT
CCTAACCTCCTCTCCCATTCAACCTGTGTAACAGGACCCAAAGGACCTGCCTCCCCGGA
AGTGCCTAACCTAGAGGGTCGGGAGGAGGTTGTGACTGACTCAGGCTGCTCCTCT
AGTTCCCTCTCATCTGACCTAGTTGCTGCCATCAGTCTAGTGGTTCGTGGTTGTC
TATTATTAACAAAAATTTGAGAACAAAAAAAAAAAAAAA

180/237

FIGURE 177

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737
><subunit 1 of 1, 1141 aa, 1 stop
><MW: 124671, pI: 5.82, NX(S/T): 5
MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDMGALRKEGEPGSLFGFSVALHRQL
QPRPQSLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWKFCFG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELSFGVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLIVGAPYFFERQEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDLNQDGFPDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGQNQYPDLLVGSILADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHGSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTWVLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPLRRLRQAPGQGLPPVAP
IILNAHQPSTQRAEIHFQKQGCEDKICQSNLQLVHARFCTRVDTEFQPLPMVDGTTALFA
LSGQPVIGLELMVTNLPDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFYLILSTSGISIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNQGQSLRTLGS AFLNIM
WPHEIANGKWLLYPMQVELEGGQGPQKGLCSPRPNILHLDVDSRDRRRRELEPP EQEPGE
RQEPMSSWWPVSAEKKKNITLDCARGTANCVVFSCPLYSFDRAAVLHVWGRLNSTFLEEY
SAVKSLEVI VRANITVKSSIKNLMLRDASTVIPVMVYLDPMMAVVAEGVPWWVILLAVLAGLL
VLALLVLLWKMFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGS PRREGP
DAHPIILAADGHPELGPDPGHPGP GTA
```

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1039-1064

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

181/237

FIGURE 178

CGCGCCGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGCGGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTCAGCAGGAGCTGCGAGCACAGTGCTGGCT
CACAAACAAGATGCTCAAGGTGTCAGCGTACTGTGTGTGCAGCCGTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGGCGGTGGCTGCAGCCGGGGGGCGGTGGACGGCGGTAAATTTC
TGGATGATAAACAAATGGCTACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCCGAGACGAAGTAGAGGATGATTATTCGCACTTGGAGTCCAGGAAAACCCTCGA
TCAGGCTTAGATCCAGCTAAGGATCCATGCTAAAGATGAAATGTAGTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAACAGGAGTAGACCATAGGCAGTGGAGGGTCCCATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCCAGCCCTGTTGTGGTTCAGATGGTCATACCTACTCTTCAGTGCA
AACTAGAATATCAGGCATGTGCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCCTCAGATAAGCCCACCAAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTCATGAAAGTGGAA
GTCAAAACAAGAACAAAAACATTGCTGAGGCCTGAGAGAACAGATTGATACCCAGC
TTGCCAATTGCAAGGACTCACTGGCTGGATGTTAACAGACTTGATAACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAACAGATTACCTTGATAAGAACAGTGTACCAAGG
CATTCTCAATTCTGTGACACATACAAGGACAGTTAACATCTAATAATGAGTGGTGC
TGCTTCCAGAGACAGCAAGACCCACCTGCCAGACTGAGCTCAGCAATATTCAAAGCGGC
AGGGGTAAAGAACCTCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAAATGTCATGGCAGTGTGGACAGTGCTGGTGTGACAGATATGAAATGAAGTC
ATGGGATCCAGAACATAATGGTGTGAGATTGTGCTATAGATTTGAGATCTCCGGAGATT
TGCTAGTGGCGATTTCATGAATGGACTGATGATGAGGATGATGAAGAACAGATATTGAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGTGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAAATTCTACATTCTAATATT
CAAAAATGATAGCCTATTAAAATTATCTTCTTCCCAATAACAAAATGATTCTAAACCTCA
CATATATTGTATAATTATTGAAAAATTGCAGCTAAAGTTAGAACCTTATGTTAAAT
AAGAACATTTGCTTGAGTTTATATTCTTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAGTTGAAGTGCTACTATAATAAAATTTCACGAGAACAAACTTGT
AAATCTCCATAAGCAAAATGACAGCTAGTGCTGGGATCGTACATGTTAATTGAAAG
ATAATTCTAAGTGAATTAAAATAAAATTGAAAGTAGCATCTGCTAGACCTAGATGAGTCAG
AAAAATATGCATGCTTAATTGCATTCAAAGTAGCATCTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAA

182/237

FIGURE 179

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVSAVLCVCAAAWCSQSLAAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFDQALDPAKDPCLMKCSRHKVCIAQDSQTAVCISHRRLTHRMRKEA
GVDHRQWRGPILSTCKQCPVVPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRSNVKRACSDLEFREVANRLRDWFKAHLHESGSQNKKTKTLLRPERSRFDT
SILPI
CKDSLGMFNRLDTNYDLLLQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCDEDGYYKPTQCHGSVGQCWCVDRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDDEDDIMNDEDEIEDDDDEEGDDDDGGDDHDVYI
```

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

183/237

FIGURE 180A

CAGACTCCAGATTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCCGGAGCAGACAA
 CAGTACCTGACGCCCTTTCAAGCCCAGGATCGCCCCAGCAGGGGATGGCGACAAGATCTGGC
TGCCCTTCCCCGTGCTCCTCTGCCGCTCTGCCCTCCGGTGTGCCTGGCGCCGGC
TTCACACCTTCCCTCGATAGCGACTTCACCTTACCCCTCCGGCAGAAGGAGTGTCT
CTACCGCCCATGCCCTGAAGGCCCGTGGAGATCGAGTACCAAGTTAGATGGAGCAG
GATTAGATATTGATTCCATCTTGCTCTCCAGAAGGCAAACCTTAGTTGAACAAAGA
AAATCAGATGGAGTTCACACTGTAGAGACTGAAGTGGTATTACATGTTCTGCTTGACAA
TACATTAGCACCATTCTGAGAAGGTGATTTCTTGAATTAACTCTGGATAATATGGAG
AACAGGCACAAGAACAGAAGATTGGAAGAAAATATTACTGGCACAGATATATTGGATATG
AAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCAGACTAAGCAAAGTGGCA
CATACAAATTCTGCTTAGAGCATTGAAGCTCGTGTGAAACATACAAGAAAGCAACTTG
ATAGAGTCAATTCTGGTCTATGTTAATTAGTGGTATGGTGGTGTGTCAGCCATTCAA
GTGTTATATGCTGAAGAGTCTGTTGAAGATAAGAGGAAAAGTAGAACTTAAAACTCCAAACT
AGAGTACGTAACATTGAAAAATGAGGCATAAAATGCAATAAAACTGTTACAGTCAAGACC
TAATGGTCTTCTCCAAATATTGAGATATAAAAGTAGGAAACAGGTATAATTAAATGTG
AAAATTAAAGTCTCACTTCTGTGCAAGTAATCCTGCTGATCCAGTTGACTTAAGTGTGTA
ACAGGAATATTGCAAGAATATAGGTTAAGTGAAGGCCATATTAAATACTGCATTTTC
CTAACTTGAAAAATTGCAATGTCTAGGTGATTAAATAATGAGTATTGGGCTAAT
TGCAACACCCAGTCTGTTTAACAGGTTCTATTACCCAGAACTTTTGTAATGCGGCAGT
TACAAATTAACTGTGAAAGTTTCAAGTTAAGTTATAAATCACCTGAGAATTACCTAATGA
TGGATTGAATAAAATCTTAGACTACAAAGCCAACTTTCTCTATTACATATGCATCTCT
CCTATAATGTAATAGAATAATAGCTTGAAGATAACATTAGGTTTTGAGATTTTATAACC
AAATACATTCACTGTAACATATTAGCAGAAAGCATTAGTCTTGTACTTGCTTACATTCC
CAAAAGCTGACATTTCACGATTCTAAAAACACAAAGTTACACTACTAAATTAGGACAT
GTGTTCTTGTAAATGAAGAATATAGTTAAAGCTTCCCTCCATAGGGACACATTTC
TCTAACCTTAACATAAAAGTGTAGGATTAAATTAATGAGGTTAAATAAGTTATT
TAATAGTATCTGTAAGTTAATATCTGTAACAGTTAATAATCATGTTATGTAATTAAAC
ATGATTGCTGACTGGATAATTCAATTACAGCAGTTATGAAGGAAATTGCTAAATG
ATCTGGCCTACCATAAAATATCTCCTTCTGAGCTAAGAATTATCAGAAAACAGG
AAAGAATTAGAAAAACTTGAGAAACCTAATCCAAAATAAAATTCACTTAAGTAGAACTAT
AAATAAATATCTAGAATCTGACTGGCTCATCATGACATCCTACTCATAACATAATCAAAGG
AGATGATTAAATTCCAGTTAGCTGGAAGAAACTTTGGCTGTAGGTTTATTCTACAAGA
ATTCTGGTTGAATTATTGTAAGCAGGTACATTATAAAATGTAAGGCCACTGTAAAG
GTTTAGCACTGGGTGTACATTATAAAATTGTTATTATAACAACATTATTAAATGG
CCTTCTGAACACTTTATTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTT
TTAAACACCTAAATGTGAATAACCCATATACAAACAAAGTTCTGCCATCTAGTTTG
AGTCTATGGGGCTTACTCAAGTACTAGTAATTAACTTCATCATGAATGAACATATAATT
TTAAGTTATGCCATTATAACGTTATGACTACATTGTGAGTTAGAAACAAACTAAA
ATTGGGGTATAGAACCCCTAACAGGTTAGTAATGCTGGAATTCTGATGAGCAATAATGA
TAACCAGAGAGTGTATTCACTCATAGTAGTATAAAAGAGATAACATTCCCTCTTA
GGCCCTGGAGAAGAGCAGCTAGATTCCCTACTGGCAAGGTTTAAATGAGGTTAA
TGCCGTATGATCAATTACCTTAATTGCCAAGAAAATGCTCAGGTGTCTAGGGGTATCC
TCTGCAACACTTGCAGAACAAAGGTCAATAAGATCCTGCCATTGAATACCCCTCCCTT
CGCTGTTAAATTGCAATGAGAACAAATTACAGTACCTAACTAATAAAGCAGGGTACAG
ATATAAAACTACTGCATTTCTATAAAACTGTGATTAAGAATTCTACCTCTCTGTATGGC
TGTTACTGTACTGTACTCTGACTCCTACCTAACATGAATTGTTACATAATCTCTAC
ATGTATGATTGTCACGTTAACCTATGATTCACTTAACCTTACCATATAAAAC
GATAATTGCTTATTGAAAAGAATTAGGAATACTAAGGACAATTATTTATAGACAAA
GTAAAAAGACAGATATTAAAGAGGCATAACCAAAAGCAAACACTGTAAACAGAGTAAAAA
TCTTTAATATTCTAAAGACATACTGTTATCTGCTTCATATGCTTTTTAATTCACTAT
TCCATTCTAAATTAAAGTTATGCTAAATTGAGTAAGCTGTTATCACTAACAGCTCATT
TGTCTTTCAATATACAAATTAAAGAATTACTACAATATTAACTAACAGGCCAACCGATT
CCATAATGTAGCAGTTACCGTGTTCACCTCACACTAACAGGCCAGGTTGCTCTGATATGCA
TTGGATGATTAATGTTATGCTGTTCTTCATGTGAATGTCAAGACATGGAGGGTGTGTA

184/237

FIGURE 180B

ATTTTATGGTAAAATTAAATCCTTCTTACACATAATGGTGTCTAAAATTGACAAAAATGAG
CACTTACAATTGTATGTCTCCTCAAATGAAGATTCTTATGTGAAATTAAAAGACATTGA
TTCCGCATGTAAGGATTTCATCTGAAGTACAATAATGCACAATCAGTGTGCTCAAACGT
CTTTATACTTATAAACAGCCATCTAAATAAGCAACGTATTGTGAGTACTGATATGTATATA
ATAAAAATTATCAAAGGAAAA

185/237

FIGURE 181

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLFPVLLAALPPVLLPGAAAGFTPSLDSDFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCFDNTFSTISEKVIFFEL
ILDNMGEQAQEQQEDWKKYITGTDILDMLKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVVVSAIQVYMLKSLFEDKRKSRT
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

186/237

FIGURE 182

CCATCCCTGAGATTTTATAAAAAACCACTTGTGACCAGACAAAGCATACCAGAT
CTCACCAAGAGAGTCGCAGACACTATGCTGCCTCCATGGCCTGCCAGTGTGCCTGGATG
CTGCTTCCTGCCTCATTCTCCTGTGTCAAGGTTAAGGTGAAGAAACCCAGAAGGAACGTGCC
CTCTCCACGGATCAGCTGTCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTGT
TTTGTCAACAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAACGGCCCTGGAAAAA
CTGGTGTCTGTGCTCAGTGGGCTGAGGGATCCTCGTGTCCCTGGTGAAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGCTCCATGACCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGAGTGGAGTAGCACTGATGTGATGAATTACTTGCATGGAGAAAATCCC
TCCACCATCTAAACCTGGCCACTGTGGAGCCTGTCAAGAACGACAGGATTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCTATGTCTGCAAGTTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTGGCGTGCAGCTCATGGACATGAGACCAGTGTGAAGAC
TCACCCCTGGAAGAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTATGATCCTCC
TTCTTTCTTTCTTCACCTCATTCAAGGCTTTCTGTCTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAAATGTTACTTATAAAAAAAAAAAAAAA

187/237

FIGURE 183

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965
<subunit 1 of 1, 175 aa, 1 stop
<MW: 19330, pI: 7.25, NX(S/T): 1
MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLWKDYNCDAKLPYVCKFKD
```

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

188/237

FIGURE 184

CCAGTCTGTCGCCACCTCACTGGTGTCTGCTGCCAGGCAAGCCTGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGACGCCGGCTGGCGCTCCTGGCGCTGGTGTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCAACCATGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCAGGAGATAGTGTACCCCTCAGGGGACTCCACGGTACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATGCCAGACCCCTGCCGTGTCTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCGCTCTGAACAGCCTCCACTGCCGGGCCCTCACGCTC
CTCCCCACTCTTGAGCCTCCGACTGTAGGTCCCCGCCACCCCCATGCCCTATGCCGCCA
GCCCGAATGCCTGAAGAAGTGCCCTGCACCAGGAAAAAAAAAAAAAAA

189/237

FIGURE 185

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405
<subunit 1 of 1, 125 aa, 1 stop
<MW: 13115, pI: 5.90, NX(S/T): 1
MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTLYSREIVYP
FQGDSTVTKSCASKCKPSDVDGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSRL
```

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

190/237

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTCTGTGGGACCCAGGCTTGC~~AA~~AGTGACGGTCATTTCCTTTCTCCCTCTT
GAGTCCTCTGAGATGATGGCTCTGGCGCAGCAGGAGCTACCCGGGTCTTGTGCGATGG
TAGCGGCGGCTCTCGGCACCCTCTGCTGGAGTGAGGCCACCTGAACTCGGTCTC
AATTCCAACGCTATCAAGAACCTGCCCCCACCGCTGGCGCTGCGGGCACCCAGGCTC
TGCAGTCAGCGCCGCCGGGAATCCTGTACCCGGCGGGATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGGAGGGACGCAGGCGTGCAAATCTGTCGCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTACGCTATGTGCTGCCCGGGATTACTGCAAAATGGAATATGTGTCTT
CTGATCAAAATCATTCGGAGGAGAAATTGAGGAAACCACACTGAAAGCTTGGTAATGAT
CATAGCACCTGGATGGTATTCCAGAAGAACCAACCTTGTCTCAAAATGTATCACACCAA
AGGACAAGAAGTTCTGTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGAAACCTGTCCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTCTAGGCTTCACACTGTCAGA
GACACTAAACCAGCTATCAAATGCAGTGAACCTTTATATAATAGATGCTATGAAAACC
TTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTCTGTTAGC
TCCAATAACACCTTCCAAAAACCTGGAGTGTAAAGAGCTTGTCTTATGGAACCTCCCTG
TGATTGCAGTAAATTACTGTATTGAAATTCTCAGTGTGGCACTTACCTGTAATGCAATGA
AACTTTAATTATTTCTAAAGGTGCTGCACTGCCTATTTCTCTGTTATGTAATTT
TTGTACACATTGATTGTTCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTAAAAGCATAACCCCTTACCCATTAAATTCTAGAGTCTAGAACGCA
AGGATCTTGGAAATGACAAATGATAGGTACCTAAAATGTAACATGAAAATAGCTTATT
TTCTGAAATGTAATCTTAAATTATTTCCCTTCTGTTAGGCTGTGATAGTTTGAA
AATAAAATTAAACATTAAAAAAAAAA

191/237

FIGURE 187

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530
<subunit 1 of 1, 266 aa, 1 stop
<MW: 28672, pI: 8.85, NX(S/T): 1
MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHFREGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCCEGLSCRIQ
KDHHQASNSSLHTCQRH
```

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

192/237

FIGURE 188

TGTGTTCCCTGCAGTCAGAATTGGGACNGCAGGGTTCCCGACCTGATTTGCAGCGGA
ACGGGAAGGTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTCTTCCTTCNG
GAGTCCTNTGAGANGATGGTTGGCGCAGCGGGAGCTAACCGGTTTGTCGATG
GTAGCGCGGTTTCGGCGGCCACCTNTGCTGGAGTGAGGCCACCTGAATCGGTTTC
AATTCCAACGNTATCAAGAACCTGCCCCACCGNTGGCGCGCTGCGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGAAATCCTGTACCCGGCGGGATAAGTACCAAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCAGACTGATGAGTACTGCGCTAGT
CCCACCCGGAGGGANGCGGGCGTCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGAAATTACTGAAAAATGGAATATGTGTNTT
CTGATAAAATCATTCGAGGAGAAATTGAGGAAACCACACTGAAAGCTTGGTAATGAT
CATAGCACCTGGATGG

193/237

FIGURE 189A

GAGGAACCTACCGGTACCGGCCGCGCTGGTAGTCGCCGGTGTGGCTGCACCTCACCAATC
 CCGTGCGCCGCGCTGGCCGTCGGAGAGTGCCTGCTCTCCTGCACGCCGGTGTGG
 GCTCGGCCAGGCCGGTCCGCCAGGGTTGAGGATGGGGAGTAGCTACAGGAAGCGAC
 CCCGCAGATGGCAAGGTATATTTGTGGAATGAAAAGGAAGTATTAGAAATGAGCTGAAGAC
 CATTCACAGATTAATATTTGGGGACAGATTGTGATGCTGATTCAACCTGAAGTAATG
 TAGACAGAAGTTCTCAAATTGCAATTACATCAACTGGAACCAGCAGTGAATCTTAATGTT
 CACTTAAATCAGAACTTGCATAAGAAAGAGAATGGGAGTCTGGTAAATAAGATGACTATA
 TCAGAGACTTGAAGGATCATTCTCTGTTCTGATAGTGTATATGCCATTAGTGGC
 ACAGATCAGGATTTACAGTTACTGGAGTGTCCAAAAGTGCAGCAGTAGAGAAATAAG
 ACAAGCTTCAAGAAATTGGCATTGAAGTTACATCCTGATAAAAACCGAATAACCCAAATG
 CACATGGCGATTTAAAAATAAGAGCATATGAAGTACTCAAAGATGAAGATCTACGG
 AAAAGATATGACAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGGCCAGTATGAAAG
 CTGGAACATTATCGTTATGATTGGTATTATGATGATGATCCTGAAATCATAACATTGG
 AAAGAAGAGAATTGATGCTGCTGTTAATTCTGGAGAACTGTGGTTGTAATTTACTCC
 CCAGGCCTGTTCACACTGCCATGATTAGCTCCCACATGGAGAGACTTGCTAAAGAAGTGG
 TGGGTTACTTCGAATTGGAGCTGTTACTGTGGTATGATAGAATGCTTGGCAATGAAAG
 GAGTCACAGCTATCCAGTCTCTCATTTCGGTCTGGAATGGCCCAAGTGAATATCAT
 GGAGACAGATCAAAGGAGAGTTAGTGAGTTTGCAATGCAGCATGTTAGAAGTACAGTGAC
 AGAACTTGGACAGGAAATTGTCACACTCCATACAAACTGCTTGTGCTGGTATTGGCT
 GGCTGATCACTTTGTCAAAAGGAGGAGATTGTTGACTTCACAGACACGACTCAGGCTT
 AGTGGCATGTTCTCAACTCATTGGATGCTAAAGAAATATATTGGAAGTAATAACATAA
 TCTTCCAGATTGAACTACTTCGGCAAACACACTAGAGGATCGTTGGCTCATCATCGGT
 GGCTGTTATTTTCATTGGAAAAAAATGAAAATTCAAATGATCCTGAGCTGAAAAAAACTA
 AAAACTCTACTTAAAAATGATCATATTCAAGTTGGCAGGTTGACTGTTCTCTGCACCCAGA
 CATCTGAGTAATCTGTATGTTTCAGCCGTCTAGCAGTATTAAAGGACAAGGAACCA
 AAGAATATGAAATTCATCATGGAAAGAAGATTCTATATGATATACTTGCCTTGCACAAAGAA
 AGTGTGAATTCTCATGTTACACGCTTGGACCTCAAAATTTCCTGCCAATGACAAAGAAC
 ATGGCTGTTGATTCTTGCCCCCTGGTGTCCACCATGTCGAGCTTACTACCAGAGTTAC
 GAAGAGCATCAAATCTCTTATGGTCAGCTTAAGTTGGTACACTAGATTGTACAGTTCAT
 GAGGGACTCTGTAACATGTATAACATTCAAGGTTATCCAACAAACAGTGGTATTCAACAGTC
 CAACATTCATGAGTATGAAGGACATCACTGCTGAACAAATCTGGAGTTCATAGAGGATC
 TTATGAATCCTTCAGTGGTCTCCCTACACCCACACCTCAACGAACAGTTACACAAAGA
 AAACACAAACGAAGTCTGGATGGTATTCTATTCTCCGGTGTATCCTGCCAAGTCTT
 AATGCCAGAATGGAAAAGAATGGCCGGACATTAACTGGACTGATCAACGTGGCAGTATAG
 ATTGCCAACAGTATCATCTTTGTCAGGAAACGTCAAAGATAACCTGAGATAAGA
 TTTTTCCCCAAAATCAAATAAGCTTACAGTATCACAGTTACAATGGTGGAAATAGGGA
 TGCTTATTCCCTGAGAATCTGGGCTAGGATTTCACCTCAAGTATCCACAGATCTAACAC
 CTCAGACTTCAGTAAAAAGTTCTACAAGGGAAAATCATTGGTGATTGATTCTATGCT
 CCTTGGTGTGGACCTGCCAGAATTGCTCCAGAATTGAGCTCTGGCTAGGATGATTAA
 AGGAAAAGTGAAGCTGGAAAAGTAGACTGTCAGGCTTATGCTCAGACATGCCAGAAAGCTG
 GGATCAGGGCCTATCCAACACTGTTAAGTTATTCTACGAAAGAGCAAAGAGAAATTTC
 GAAGAGCAGATAAATACCAAGAGATGCAAAAGCAATCGTCGCCATAAAAGTGAAGA
 AACTCTCCGAAATCAAGGCAAGAGGATAAGGATGAACTTGTATAATGTTGAAGATGAAGAA
 AAAGTTAAAAGAAATTCTGACAGATGACATCAGAAAGACACCTATTAGAATGTTACATT
 TGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGACTGCCAGAATTATCTACAGCA
 CTGGTGTAAAAGAAGGGCTGCAAACCTTTCTGTAAGGGCCGGTTATAAATATTTAGA
 CTTTGCAGGCTATAATATGGTCACACATGAGAACAGAAATAGAGTCATCATGTATTCTT
 TGTTATTGCTTTAACACCTTAAAAAATATTAAACGATTCTAGCTCAGAGCCATACA
 AAAGTAGGCTGGATTCACTGGACCATAGATTGCTGCCCTCGACGGACTTATAATG
 TTTCAAGTGGCTGGCTGAACATGAGTCTGCTGTGCTATCTACATAAATGCTAAGTGTAT
 AAAGTCCACTTCCCTCACGTTTTGGCTGACCTGAAAAGAGGTAACCTAGTTGGTC
 ACTTGTCTCCTAAAAATGCTATCCCTAACCATATATTATTCGTTTAAAAACACCCA
 TGATGTGGCACAGTAAACAAACCTGTTATGCTGTATTATGAGGAGATTCTTCATTGTT
 TTCTTCCCTCTCAAAGGTTGAAAAATGCTTTAATTTCACAGCCGAGAAACAGTCAG

194/237

FIGURE 189B

CAGTATATGTGCACACAGTAAGTACACAAATTGAGCAACAGTAAGTGCACAAATTCTGTAG
TTTGTCTATCATCCAGGAAAACCTGAGGGAAAAAATTATAGCAATTAACTGGGCATTGTA
GAGTATCCTAAATATGTTATCAAGTATTAGAGTTCTATATTAAAGATATATGTGTTCAT
GTATTTCTGAAATTGCTTCATAGAAATTCCCCTGACTGATAGTTGATTGAGGCATCTA
ATATTTACATATTCGCCTCTGAACCTTGTGACCTGTATCCTTATTTACATTGGGTTT
TTCTTCATAGTTGGTTTCACTCCTGTCCAGTCTATTATTCAAATAGGAAAAT
TACCTTACAGGTGTTTACTGTAGCTATAATGATACTGTAGTTATCCAGTTACTAGTT
ACTGTCAAGGGCTGCCTTTCAAGATAAAATTGACATAAACTGAAGTTATTTTATAAA
GAAAATCAAGTATATAAATCTAGGAAAGGGATCTCTAGTTCTGTGTTAGACTCAAA
GAATCACAAATTGTCAGAACATGTAGTTAGTTATAATTCAAGGTACAGAATGGT
AAAAATTCCAATCAGTCAAAAGAGGTCAATGAATTAAAAGGCTTGCAACTTTCAAAAAAA
AAAAAA

195/237

FIGURE 190

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTDQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYYRYDFGI
YDDDPEIITLERREFDAAVNSGELWFVNFYSPGCSHCHDLAPTRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLKNDHIQVGRFDCSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIHGKKILYDILAFAKESVNSHVTTLGPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVFNQSNIHEYEGHHS
AEQILEFIEDLMNPSSVSLPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWRNDAYSRLRIWGLG
FLPQVSTDLPQTSEKVLQGKNHWVIDFYAPWCGPCQNFAPEFELLARMIKGKVAGKVDC
QAYAQCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAIAAALISEKLETLRNQGKRNKDEL
```

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

196/237

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTCTGCTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTCATTCCCTCAGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTGCAAAACGACAGAGC
ATATTGGTTCTGTGGATATTAATAAGCGCGGTGTGGAGGAAACTGCAGCTGAGTGCCGAAA
ACTAGGCCTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTCAGCACCAAGGATGAAGAGATTACCAAGACATTGAGGTCAA
CATCCTAGGACATTTGGATCACAAAGCACTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCTACCTCATCCCATA
TGTTCAGCAAATTGCGCTGTGGCTTCACAGAGGTCTGACATCAGAACTTCAGGCC
GGGAAAAACTGGTATCAAAACCTCATGTCCTGCCAGTTTGTGAATACTGGGTCACCA
AAAATCCAAGCACAAGATTATGCCCTGTATTGGAGACAGATGAAGTCGAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTGTTCCATCGTATATCAATATCTTCTGAG
ACTACAGAAGTTCTTCTGAACGCGCCTCAGCGATTAAATCGTATGCAGAAATTCAAT
TTGAAGCAGTGGTGGCCACAAATCAAATGAATAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTGAATCAATGCTGCAAAGCTTATTCACATTT
TCAGTCCTGATAATATTAAAAACATTGGTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCTGTTCTCAAGAATATTACGTAGTTTCAAGGTCTGTTTCT
TCATGCCTCTTAAAAACTTCTGTGCTTACATAAACATACTTAAAGGTTTCTTAAGATAT
TTTATTTCATTTAAAGGTGGACAAAGCTACCTCCCTAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTAAGACTGTTCAAGTAGCATTCAATCTGTAGCCATGCCACAG
AAATATCAACAAGAACACAGAACATGAGTGCACAGCTAAGAGATCAAGTTCAAGCAGG
ATCTCAACCTGGACATATTAAAGATTCAAGCATTGAAAGATTCCCTAGCCTCTCCTT
TCATTAGCCAAAACGGTGCAACTCTATTCTGGACTTATTACTGATTCTGTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCCTCTATATTCCCTCCCTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTAAATCTCAGAATTAAAGTTCTAGCCCCATGA
TAACCTTTCTTGTAAATTATGCTTCAATATCCTGGTCCCAGAGATGTTAGACAAT
TTTAGGCTAAAAATTAAAGCTAACACAGGAAAGGAACGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

197/237

FIGURE 192

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409
<subunit 1 of 1, 300 aa, 1 stop
<MW: 33655, pI: 9.31, NX(S/T): 1
MNIILEILLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSI
LVLWDINKRGVEETAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTV
YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCVGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLPVFVNTGFTKNPSTRLWPVLETDEVVRSLID
GILTNKKMIFVPSYINIFRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK
```

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein
amino acids 165-202, 37-49, 112-122 and 210-219

198/237

FIGURE 193

CGGCGGCGGCTCGGGCGCGAGGTGAGGGCGCGAGGTGAGGGCGCGAGGTCCAGCAGG
ATGCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCGGAGAGGGCCAGCCCAGGGC
AGGATGACCAAGGCCCCGGCTGTTCCGGCTGTGGCTGGTCTGGGTGGTGTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCCGCGCACCTCTACTTGCACACGTCTTCT
CTAGGCCGCACACGGGCCGCGCTGCCACGCCGGGCGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTCTGGACAAGTTCTCAGTGCTGGCGTAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCAGCCGGGAGCATGGAGGAGAGCGTGA
GCTACGACTGGTCCCCGCGACGCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCCGAG
CGGAGGAGCGTGCCTGCAGGGCTTCTGCAGCCAACCTCCAGCCTGGCCTCCCCACCAAGGAGCG
CGCATTGACGACATCCCCAACCTGGAGCTGAGCCACCTGATCGTGGACGACGGCACGGG
CCATCTACTGCTACGTGCCAACGGTGGCCTGCACCAACTGGAAGCGCGTGTGATGATCGTGT
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGAGCA
CGTGCACAACGCCAGCGCGCACCTGACCTCAACAAGTTCTGGCGCCCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCTCGTGCACGCC
TTCGTGCGCCTGATCTCCGCCTCCGCAGCAAGTTGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTGCCGTGCCCATGCTGCCGTGACGCCAACACACCAGCCTGCCGCCCTGGCG
GCGAGGCCTCCCGCTGCCCTCAAGGTGTCCTCGCCAACCTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGCGCCCTCAACGAGCACTGGCGCAGGTGTACCGCCTCTGCCA
CCCGTGCCAGATCGACTACGACTTCGTGGGAAGCTGGAGACTCTGGACGAGGACGCC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTCCCCCGAGCTACCGAACAGG
ACCGCCAGCAGCTGGAGGAGGACTGGTGCACAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAAACTCTACGAGGCCGACTTTGTTCTTCGGCTACCCCAAGCCGAAAACCTCCTCC
GAGACTGAAAGCTTCGCGTTGCTTTCTCGCGTGCCTGGAACCTGACGCACGCCACTCC
AGTTTTTATGACCTACGATTTGCAATCTGGCTTCTGTTCACTCCACTGCCCTATACCC
ATTGAGTACTGTATCGATATTGTTTAAGATTAATATTCAAGGTATTAATACGA

199/237

FIGURE 194

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112
<subunit 1 of 1, 414 aa, 1 stop
<MW: 48414, pI: 9.54, NX(S/T): 4
MTKARLFRLWLVLSVFMILLIIVYWDAGAAHFLHTFSRPHTGPLLPTPGPDRLDRELTA
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSILAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLS
GSLLHRGAPYRDPRLIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKKYTKFLFVRDPF
VRLISAFRSKFELENEEFYRKFAVPMRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDDEAAQLLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQQLYKLYEADFVLFGYPKPENLLRD
```

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

200/237

FIGURE 195

TCGGGCCAGAATTGGCACGAGGCCACGAGGGGACGGCCTCACGGGCTTGGAGGTGA
AAGAGGCCAGAGTAGAGAGAGAGAGACCGACGTACACGGATGGCACGGAACCGCCT
ATGCCGGGAAGGTGGTGGTCGTGACCGGGGCGGGCGGCATCGGAGCTGGGATCGTGC
GCCTTCGTGAACAGCAGGGCCCGAGTGGTTATCTCGACAAGGATGAGTCTGGGGCCGG
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCCTGGTTCTGAGACCATCCGCCGATTGGCCGCTGGATTGTGTTCAACAAAC
GCTGGCCACCACCCACCCCCACAGAGGCCCTGAGGAGACCTCTGCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGACGTACACCTTGACCAAGCTGCCCTCCCACCTGCCGA
AGAGTCAAGGAATGTCATCAACATCTCCAGCCTGGTGGGGCAATCGGCAGGCCAGGCA
GTTCCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCGCTGTGG
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCGCTGTGG
AGGAGCTGGCAGCCTTAATGCCAGACCTAGGGCCACAATCCGAGAGGGCATGCTGGCCAG
CCACTGGGCCGCATGGCCAGCCGCTGAGGTGGGGCTGCCAGTGTCCCTGGCCTCCGA
AGCCAACCTCTGCACGGCATTGAACGTGCTGACGGGGGTGCAGAGCTGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGGACGCCCGATATCCCTGTGATTCTCTCATT
CTACTGGGCCCCCTCCTAGGACTCTCCACCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTAGACTCTAACGCCAGTTAGCAAGGTGCCGGTCACCCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

201/237

FIGURE 196

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVTGGGRGIGAGIVRAFVNNSARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRRLDCVVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLTKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDES PYGVRVNCISPGN
IWTPWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS
```

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

202/237

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAAACCTG
CTGTTCTTCTTACCATTCATCTCCTGGGCTGGGCCAGCCCAGGAGCCCCAAAGCAA
GAGGAAGGGCAAGGGCGGCCTGGGCCCTGGCCCTGGCCCTACCAGGTGCCACTGGACC
TGGTGTACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCAGAGAAAGTGTGAGGTCAACTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGCTGTGAACCCCTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCGGTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCGCCACCGCCCCGCACAGGGCCTGCCGCCAGCGCGCAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTCTGAATCACCTGGCCAGAAGCCAGGCCAGCAGCCGAGA
CCATCCTCCTTGCACCTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTGAAA
GCAAG

203/237

FIGURE 198

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294
<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pi: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVSDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPRTGPCRQRAVMETIAVGCTCIF
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

204/237

FIGURE 199

GCGCCGCCAGGCGTAGGC GG GTGGCCCTT GCGT CCCC GCTT GAAAAACCCGGCGGG
CGAGCGAGGCTGCGGGCCGGCCGCTGCCCTCCCCACACTCCCCGCCAGAAGCCTCGCTCG
GCGCCAACATGGCGGTGGCGCTCGGGCCCGCAGCTAACGGCGCTCTGGCCGCTGGAT
CGCGCGTGTGGCGGCAGGGCAGGGCCGAGGGCCGCTGCCGCCGGAGCAGAGCCGG
TCCAGCCCATGACCGCCTCAACTGGACGCTGGTATGGAGGGCGAGTGGATGCTGAAATT
TACGCCCATGGTGTCCATCCTGCCAGCAGACTGATTCAAATGGGAGGGCTTTGCAAAGAA
TGGTCAAATACTTCAGATCAGTGTGGGGAAAGGTAGATGTCATTCAAAGAACAGGTTGAGTG
GCCGCTCTTGTCAACCCTCTCCAGCATTTCATGCAAAGGATGGATATTCCGCCGT
TATCGTGGCCAGGAATCTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAATCCCCAGCTCTCTAACGATGTCCTGGAATGGCTGGTC
TTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACATTTCACAGTGACTCTTGGAAATT
CCTGCTTGGTGTCTTATGTGTTTCGTCAAGCCACCTGGTTTGGCCTTTATGGG
TCTGGTCTTGGGGTAATATCAGAATGTTCTATGTGCCACTTCAAGGCATTATCTGAGC
GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCAGGAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAACAAAGACAGCCTGTAGATGATGAAGAAGA
GAAAGAAGATCTGGCGATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGACAACCTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCATGATCAGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTGAGCAACCCCTGCCAGC
TGACACAGAGGTGGAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTAATGATGCTTTCAAAGATAACACACAAATATGTCAGCTCCCTTGG
CCTGCAGTTGTACCAAATCCTAATTTCCTGAATGAGCAAGCTCTCTTAAAGATGCT
CTCTAGTCATTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGATCTGTTGGAGACTGGATGGAA
CAAGTTCAATTACTTAGGGTCAGAGAGTCGACCAAGAGGAGGCCATTCCCAGTCCTAATC
AGCACCTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTGGCT
CCTGAGCATCCCCAAAGTGTAAAGCTAGAAGCCTGCATCCTTCTGTGAAAGTATTAT
TTTGTCAAATTGCAAGGAAACATCAGGCACCACAGTCAGTAAAAAATCTTCACAGCTAGAA
ATTGAAAGGGCTTGGTATAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTATTCTTACCTTAATTTCAGCATTCCACCAGGGCATTAGGCTCT
CCACACTCTCACTATTATCTCTGGTCAGAGGACTCCAATAACAGCCAGGTTACATGAAC
TGTGTTGTCATTGACCTAAGGGTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGTCAGAGACTCAAAGGAAGTAAGGATT
TTACAAGACAGATTTAAAGGTTGTTGTCAAAATATAGTTGTTGATTTTTTT
AAGTTTCTAAGCAATATTTCAAGCCAGAAGTCCTCTAAGTCTGAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTGTTCATCTCAAGGGGTTCCCTGGGTCTGAAC
TACTTAATAATAACTAAAAACCACTCTGATTTCTTCAGTGTGCTTTGGTAAA
GAATTAAATGAACCTCAGTACCTGAAAGTGAAGATTGATTGTTCCATCTCTGTAATC
TCCAAAGAATTATATCTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTCTT

205/237

FIGURE 200

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRCGPQLTALLAAWIAAVAAATAGPEEAALPPEQSRVQPMTASNWTLMGEWMLKFYAP
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTLPAFFHAKDGFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDEAEEEEEDNLAAAGVDEERSEANDQGPPGEDGVTR
EVEPEEAESEGISEQPCPADTEVVVEDSLRQRKSQHADKGL
```

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

206/237

FIGURE 201

ATCTGGTTGAAC TACTTAAGCTTAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATCTCTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAAGAAAAGAGCTCC
CAAATGCTATATCTATTCAAGAACAATGGAATATCATCCTGATTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTTCAGAGAAAGGATCGTGTGCATCTCCTGGCGCCTCATGCTGTAATTTGG
GAATCCTATGCTGGTAATTA CACTGGTGTAGCTGTGGCCTGGTACCATGGGTTCTTCC
AGCCCTTGTCCCTCTAATTGGATTATGAGAAGAGCTGTATCTATTCAAGCATGTCACT
AAATTCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCTAAAGATAG
ACAGCTCAAATGAATTGGGATTATAGTAAAACAAGTGTCTCCAACCTGATAATTCAATT
TGGATAGGCCTTCTCGGCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTCAGATCAGAACACAGCTACCCAAGAAAACCCATCTCAAATTGTG
TATGGATTACGTGTAGTCATTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTG
GAGAAGAAGTTCAATGTAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAATGCAGAAAATG
TTTAGAGAGCTGGCCA ACTGTAATCTTAACCAAGAAATTGAAGGGAGAGGCTGTGATTCT
GTATTGTCGACCTACAGGTAGGCTAGTATTATTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTGAGTTTATTTTTATTATTTTGAGATAGGGTCT
CACTTTGTACCCAGGCTGGAGTGCAGTGGCACAACTCGACTCACTGCAGCTATCTCGC
CTCAGCCCCCTCAAGTAGCTGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTGGT
TTTTTGAGAGACTGGGTTTGCATGTTGACCAAGCTGGCTCTAACTCCCTGGGCTTAAG
TGATCTGCCCGCCTGGCCTCCAAAGTGTGGATTACAGATGTGAGCCACCACACCTGG
CCCAAGCTGAATTTCATTCTGCCATTGACTGGCATTACCTGGTAAGCCATAAGCGA
ATCTTAATTCTGGCTCTATCAGAGTTGTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTGACCCCTCAACTCTAGCAGTATATCAGTTAGTAAC TGAGGGTGA
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAATCTCACCACAGTCAGAGCAATT
ATTATTTCTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAGCAGGAAC
CTTCTTCTTATCAATTAAATAGCTCAGAGAGTACATGCCATATCTCAATAGAATCTT
TTTTTTTTTTTTTGAGACAGAGTTCGCTCTTGCCAGGCTGGAGTGCAACCG
CACGATCTGGCTCACCGCAACCTCCGCCCCCTGGTCAAGCAATTCTCCTGCCTCAGC
CCCAAGTAGCTGGGATTACAGTCAGGCACCACACCCGGCTAATTGTATT
AGAGACAGGGTTCTCCATGTCGGTCAGGGTAGTCCGAACCTCTGACCTCAAGTG
CTGCCTCGGCCTCCAAAGTGTGGATTACAGGCGTGAGCCACTGCACCCAGCCTAG
TGTATAATATGTAATTGTAGGGAAACTGCTCTCATAGGAAAGTTCTGCTTTAA
AAAATACATAAAATACATAAAATCTGATGATGAATATAAAAAGTAACCAAC
CTCATTGGAAACAATTTGGAATATGTTATTAGTTGTGATGTACTGTTACA
ACCATTTTCTCAGTAATTACTGTAAAATGGTATTATTGGAATGAAACT
ATATTCTCATG
TGCTGATTGTCTTATTTCATACTTCCACTGGTGTATT
TTCTGTATTACTAGGGAGGCATTACAGTCCTCTAATGTTGATT
TGTACCAATTACTAAATTATGCA GTTTAAATGGATGATT
TTGTTATGTTATGTTGAGGATTTCAT
TTCAATAAAAAAAAAACTCTTATCAAAAAAAAAAAAAAAA
207/237

FIGURE 202

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIYEKSCYLFSMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM
```

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

208/237

FIGURE 203A

GGAAGGGAGGAGCAGGCCACACAGGCACAGGCCGGTGAGGGACCTGCCAGACCTGGAGGG
 TCTCGCTCTGTACACAGGCTGGAGTCAGTGGTGTGATCTGGCTCATCGTAACCTCCACC
 TCCC GG GT TCAAGT GATT CT CAT GCCT CAG CCT CCC GAG T AGCT GGG ATT ACAGG TGG T GAC
 TTCCAAGAGTGACTCCGTGGAGGAAAATGACTCCCCAGTCGCTGCTGCAGACGACACTGTT
 CCTGCTGAGTCTGCTCTTCCCTGGTCCAAGGTGCCACGGCAGGGGCCACAGGGAAAGACTTTC
 GCTTCTGCAGCCAGCGAACAGACACACAGGAGCAGCCTCACTACAAACCCACACCAGAC
 CTGCGCATCTCCATCGAGAACTCCGAAGAGGCCCTCACAGTCCATGCCCTTCCCTGCAGC
 CCACCCCTGCTTCCCGATCCTCCCTGACCCAGGGCCCTTACCACTCTGCCTCTACTGGA
 ACCGACATGCTGGGAGATTACATCTCTATGGCAAGCGTGA CTTCTGCTGAGTGACAAA
 GCCTCTAGCCTCCTCTGCTTCCAGCACCAGGAGGAGGCCTGGCTCAGGGCCCCCGCTGTT
 AGCCACTTCTGTCA CCTCCTGGTGGAGCCCTCAGAACATCAGCCTGCCAGTGCCGCCAGCT
 TCACCTTCTCCTCCACAGTCCTCCCCACACGGCCGCTCACAAATGCCCTGGTGGACATGTGC
 GAGCTCAAAGGGACCTCAGCTGCTCAGCCAGTTCTGAAGCATCCCCAGAAGGCCTCAAG
 GAGGCCCTCGGCTGCCCGCCAGCCAGCAGCTGCAAGAGCCTGGAGTCGAAACTGACCTCTG
 TGAGATTCATGGGGGACATGGTGTCTCGAGGAGGACCGGATCAACGCCACGGTGTGGAAG
 CTCCAGCCCACAGCCGGCCTCAGGACCTGCACATCCACTCCGGCAGGGAGGAGCAGAG
 CGAGATCATGGAGTACTCGGTGCTGCTGCCCTCGAACACTCTCAGAGGAGCAGAAGGCCGGA
 GCGGGGAGGCTGAGAAGAGACTCCTCCTGGTGGACTTCAGCAGCCAAGCCCTGTTCCAGGAC
 AAGAATTCCAGCCAAGTCCTGGGTGAGAAGGTCTTGGGATTGTTGACAGAACACCAAAGT
 AGCCAACCTCAGGGAGCCGTGGTGTCACTTCCAGCACCAGCTACAGCCGAAGAATGTGA
 CTCTGCAATGTGTGTTCTGGGTTGAAGACCCCACATTGAGCAGCCGGGATTGGAGCAGT
 GCTGGGTGTGAGACCCTCAGGAGAGAAACCAAACATCCTGCTTCTGCAACCACCTGACCTA
 CTTTGCACTGCTGATGGTCTCTCGGTGGAGGTGGACGCCGTGCACAAGCACTACCTGAGCC
 TCCCTCCTACGTGGCTGTGCTCTGCCCTGCCCTGTCA CACCATTGCCGCTAC
 CTCTGCTCCAGGGTGCCCTGCCGTGCAGGAGGAAACCTCGGGACTACACCATAAGGTGCA
 CATGAACCTGCTGCTGGCGTCTCCTGCTGGACACGAGCTCCTGCTCAGCGAGCCGGTGG
 CCCTGACAGGCTCTGAGGCTGGCTGCCAGCCAGTGCATCTCCTGCACTTCTCCCTGCTC
 ACCTGCCCTTCCGGATGGGCTCGAGGGGTACAACCTCTACCGACTCGTGGTGGAGGTCTT
 TGGCACCTATGTCCCTGGCTACCTACTCAAGCTGAGCGCCATGGGCTGGGCTTCCCATCT
 TTCTGGTACGCTGGTGGCCCTGGATGTGGACA ACTATGGCCCCATCATTTGGCTGTG
 CATAGGACTCCAGAGGGCTCATCTACCCCTCCATGTGCTGGATCCGGACTCCCTGGTCAG
 CTACATCACCAACCTGGGCTCTCAGCCTGGTGTGTTCTGTTCAACATGCCATGCTAGCCA
 CCATGGGGTGCA GATCTGCGGCTGCCCTGGGACACCCAAAAGTGGTACATGTGCTGACA
 CTGCTGGGCTCAGCCTGGTCTTGGCTGCCCTGGGCTTGGATCTTCTCTCCCTTGGCTTC
 TGGCACCTCCAGCTTGTGCTCTACCTTTCA GGCATCATCACCTCCCTCAAGGCTTCC
 TCATCTCATCTGGTACTGGTCCATGCGGCTGCAGGCCGGGTTGGCCCTCCCTGTAAG
 AGCAACTCAGACAGCGCCAGGCTCCCATCAGCTGGCAGCACCTCGTCCAGCCGATCTA
GGCCTCCAGCCCACCTGCCCATGTGATGAAGCAGAGATGCCCTCGCACACTGCCTGT
GGCCCCCAGCCAGGCCAGGCCAGGCCAGTCAGCCGAGACTTGGAAAGCCCAACGACC
ATGGAGAGATGGGCGTTGCCATGGTGGACGGACTCCGGCTGGGTTTGAAATTGGCCTT
GGGGACTACTCGGCTCTCACTCAGCTCCACGGGACTCAGAAAGTGCAGGCCATGCTGCC
GGGTACTGTCCCCACATCTGCTCCAACCCAGCTGGAGGCCTGGTCTCTCCTTACAACCC
GGCCCA GCGCCCTCATTGCTGGGGCCAGGCCCTGGATCTTGAGGGCTGGCACATCCTTA
ATCCTGCTGCCCTGCCCTGGGACAGAAATGTGGTCCAGTTGCTCTGCTCGTGGTCA
CCCTGA GGGCACTCTGCATCCTCTGTCATTAACTCAGGTGGCACCCAGGGCGAATGGG
CCAGGCCAGGCCAGGCCAGGCCCTGGCGAGGAGAGGCCCTTGCCAGGAGCACAGCAG
CTCGCCTACCTCTGAGCCCAGGCCCTCCCTCCCTCAGCCCCCAGTCCTCCCTCATCT
CCCTGGGTTCTCCTCTCCAGGGCCTTGCTCCTGTTCAAGCTGGGGTCC
GATTCCAATGCTGTTTTGGGGAGTGGTTCCAGGAGCTGCCCTGGTGTCTGCTGTAATGT
TTGTCTACTGCACAAGCCTCGGCCTGCCCTGAGCCAGGCTCGGTACCGATGCGTGG
GCTAGGGCCCTGTCCATCTGGGCTTTGTATGAGCTGCATTGCCCTGCTCACCC
CTGACC AAGCACACGCCCTCAGAGGGGCCCTCAGCCCTCCTGAAGCCCTTTGTGG
CAAGAAGACTGAGACTGACCTCCACACTCCAAAGGACTGAGACTGACCTCCCTG
GTGACACTGGCCTAGAGCCTGACACTCTCTTAAGAGGTTCTCTCCAAGCCCCAA
ATAGCTC

209/237

FIGURE 203B

CAGGCGCCCTCGGCCGCCATCATGGTTATTCTGTCCAACAAACACACACGGTAGATTGC
TGGCCTGTTGAGGTGGTAGGGACACAGATGACCGACCTGGTCACTCCTGCCAACATTC
AGTCTGGTATGTGAGGCGTGCCTGAAGCAAGAACTCCTGGAGCTACAGGGACAGGGAGCCAT
CATTCCCTGCCTGGGAATCCTGGAAGACTTCCTGCAGGAGTCAGCGTCAATCTTGACCTTGA
AGATGGGAAGGATGTTCTTTACGTACCAATTCTTGTCTTGATATTAAAAAGAAGTA
CATGTTCATGTAGAGAATTGGAAACTGTAGAAGAGAATCAAGAAGAAAAATAAAAATCAG
CTGTTGTAATGCCCTAGCAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

210/237

FIGURE 204

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCQSQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKDFLSDKASSLLCFQH
QEESLAQGPPPLLATSVTSWWSPQNISLPSAASFTFSFHSPPHTAAHNASVDMCELKRDLQLL
SQFLKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLOPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGEAEKRLLVDFSSQALFQDKNSSQVLGE
KVLGIVVQNTKVANLTEPVVLTQHQHQLQPKNVTLQCVFWVEDPTLSSPGHWSSAGCETVRRE
TQTSCFCNHLYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVSALACLVTIAAYLCSRVPPLPC
RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTIVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSDLVSYITNLGLFSLVFLFNMMAMLATMVVQILRLRPTHQWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGPSPLKSNSDSARLP
ISSGSTSSSRI
```

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

211/237

FIGURE 205

TGCCTGGCTGCCTGTCAACAATGCCGTTACTCTGCTTCCAGGTTGCCCTGCCTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGCACATGAACCTGCTGCTGGCCGTCTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTCCTGCACTTCTCCTGCTCACCTGCCTTCCTGGATGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTGGCACCTATGTCCCTGGTACCTACTCAA
GCTGAGCGCCATGGCTGGGCTTCCCCATCTTCTGGTACGCTGGTGGCCCTGGTGGATG
TGGACAACATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCC
TCCATGTGCTGGATCCGGACTCCCTGGTCAGCTACATACCAACCTGGCCTTCAGCCT
GGTGGTCTGTTAACATGG

212/237

FIGURE 206

CGGACGCGTGGCGGACCGTGGCGGACCGTGGCGGACCGTGGCTGGTCAGGTCCA
GGTTTGCTTGTATCCTTTCAAAAACGGAGACACAGAAGAGGGCTCTAGGAAAAAGTTT
GGATGGGATTATGTGGAAACTACCCCTGCATTCTGCTGCCAGAGCAGGCTGGCGCTTCC
ACCCCAGTGCAGCCTTCCCCTGGCGGTGGTAAAGAGACTCGGGAGTCGCTGCTTCAAAGT
GCCCGCCGTGAGTGAGCTCTCACCCAGTCAGCAAATGAGCCTCTCGGGCTCTCCTGCT
GACATCTGCCCTGGCGGCCAGAGACAGGGACTCAGGCAGAATCCAACCTGAGTAGTAAAT
TCCAGTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAAATTATT
ACTGTGTCTACTAATGGAAGTATTACAGCCAAAGGTTCCCTCATACTTATCCAAGAAATAC
GGTCTTGGTATGGAGATTAGTAGCAGTAGAGGAAATGTATGGATACAACCTACGTTGATG
AAAGATTGGGCTTGAAGACCCAGAAGATGACATATGCAAGTATGATTTGAGAGTTGAG
GAACCCAGTGTGGAACTATATTAGGCGCTGGTGTGGTCTGGTACTGTACCAAGGAAACA
GATTCTAAAGGAAATCAAATTAGGATAAGATTGTATCTGATGAATATTTCCTCTGAAC
CAGGGTCTGCATCCACTACAACATTGTATGCCACAATTACAGAGCTGTGAGTCCTTCA
GTGCTACCCCTTCAGCTTGCCACTGGACCTGCTTAATAATGCTATAACTGCCCTTAGTAC
CTTGGAAAGACCTTATTGATATCTGAAACCAGAGAGATGGCAGTTGGACTAGAAGATCTAT
ATAGGCCAACTTGGCAACTTCTTGGCAAGGTTTGTGAGGAAATCCAGAGTGGT
GATCTGAACCTCTAACAGAGGAGTAAGATTATACAGCTGCACACCTCGTAACCTCTCAGT
GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTCTGGCCAGGTTGTCCTGGTTA
AACGCTGTGGTGGAACTGTGCCTGTTCCACAATTGCAATGAATGTCAATGTGTC
AGCAAAGTTACTAAAAAATACCAACGGAGGCTTCAGTTGAGACCAAGACCAGGTGCA
ATTGCACAAATCACTCACCAGCTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTC
GAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAGCTTGCCAGAGCTGTGCA
AGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCATCTTAATCTCAGTTGTTGCTT
CAAGGACCTTCATCTCAGGATTACAGTCATTCTGAAAGAGGAGACATCAAACAGAATT
AGGAGTTGTGCAACAGCTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTCAATC
GTGGAAAGAAAATTAAATGTTGTTAAATAGATCACCAGCTAGTTCAGAGTTACCATGTA
CGTATTCCACTAGCTGGTTCTGTATTCAGTTCTGATACGGCTTAGGGTAATGTCAGT
ACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCTTAACTCTAAAGCTCC
ATGTCTGGCCTAAATCGTATAAAATCTGGATTTTTTTTGCTCATATTCA
ATATGTAAACCAGAACATTCTATGACTACAAACCTGGTTTAAAAGGAACATATGTTGCT
ATGAATTAAACTTGTGTATGCTGATAGGACAGACTGGATTTCATATTCTTATTAAAAT
TTCTGCCATTAGAAGAAGAGAACATACATTGATGGTTGGAAGAGATAAACCTGAAAAGAAG
AGTGGCCTATCTCACTTATCGATAAGTCAGTTATTGTTGCTTGTGACATTTTTAT
ATTCTCCTTTGACATTATAACTGTTGGCTTTCTAATCTGTTAAATATATCTATTTCAC
CAAAGGTATTAAATATTCTTTTATGACAACTTAGATCAACTATTCTAGCTGGTAAATT
TTCTAAACACAATTGTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGAC
AAAAATACATGTATTCTCATCTCGTATGGCTAGAGTTAGATTAATCTGCATTAAAAA
CTGAATTGGAATAGAATTGTAAGTTGCAAGACTTTGAAAATAATTAAATTATCATATC
TTCCATTCTGTTATTGGAGATGAAAATAAAAGCAACTTATGAAAGTAGACATTCA
AGCCATTACTAACCTATCCTTTGGGAAATCTGAGCCTAGCTCAGAAAACATAAAGC
ACCTTGAAAAGACTTGGCAGCTCCTGATAAAGCGTGTGCTGCAAGTAGGAACACAT
CCTATTATTGTGATGTTGTTTATTATCTTAAACTCTGTTCCATACACTGTATAAAT
ACATGGATATTCTATGTACAGAACGTATGTCCTTAACCAAGTTCACTTATTGTA
ATTAAAAGAAAATCAGTAAATATTGCTTGTAAAATGCTTAATATNGTGCCTAGGTT
GTGGTGAATTTGAATCAAAATGTATTGAATCATCAAATAAAAGAATGTGGCTATTG
GGAGAAAATTAAAAAAAAAGGTTAGGGATAACAGGGTAATGCAGGCC

213/237

FIGURE 207

MSLFGLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRRVVDLNLLTEEVRLY
SCTPRNFVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHECDCVCRGSTGG

214/237

FIGURE 208

CCCATCTCAAGCTGATCTGGCACCTCTCATGCTCTGCTCTTCAACCAGACCTACATT
CCATTTGGAAGAAGACTAAAAATGGTTTCCAATGTGGACACTGAAGAGACAAATTCTTA
TCCTTTAACATAATCTAATTCCAAACTCCTGGGGCTAGATGGTTCTAAACTCTG
CCCTGTGATGTCACTCTGGATGTTCAAAGAACCATGTGATCGTGGACTGCACAGACAAGCA
TTTGACAGAAATTCTGGAGGTATTCCCACGAACACACCACGAACCTCACCCCTACCAATTAAACC
ACATACCAAGACATCTCCCCAGCGTCCTTCACAGACTGGACCATCTGGTAGAGAGATCGATTTC
AGATGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACACATGTGCATCAAGAGGGTGCA
GATTAACCCAGAAGCTTAGTGGACTCACTTATTAAAATCCCTTACCTGGATGGAAACC
AGCTACTAGAGATAACCGCAGGGCTCCGCCTAGCTTACAGCTCTCAGCCTGAGGCCAAC
AACATCTTCCATCAGAAAAGAGAATCTAACAGAACTGGCAACATAGAAATACTCTACCT
GGGCCAAAACTGTTATTATCGAAATCCTGTTATGTTCATATTCAATAGAGAAAGATGCCT
TCCTAAACTTGACAAAGTTAAAGTGCTCTCCCTGAAAGATAACAAATGTACAGCGTCCCT
ACTGTTTGCCATCTACTTAAACAGAACTATATCTACAAACAACATGATTGCAAAATCCA
AGAAGATGATTAAATAACCTAACCAATTACAAATTCTTGACCTAAGTGGAAATTGCCCTC
GTTGTTATAATGCCCATTCCTGTGCCGTGAAAAATAATTCTCCCTACAGATCCCT
GTAAATGCTTTGATGCCGTGACAGAATTAAAAGTTTACGTCTACACAGTAACCTCTTCA
GCATGTGCCCAAGATGGTTAAGAACATCAACAAACTCCAGGAACACTGGATCTGCCAAA
ACTTCTGGCCAAGAAATTGGGGATGCTAAATTCTGCATTTCTCCCTAGCCTCATCCAA
TTGGATCTGTCTTCAATTGAACTTCAGGTCTACGTGCATCTATGAATCTATCACAAGC
ATTTCCTCACTGAAAAGCCTGAAAATTCTGCCATTACATAATCTCAAAATCTGAAGTTCTGATCTGGCACT
AAAGCTTAAACCTCTGCCATTACATAATCTCAAAATCTGAAGTTCTGATCTGGCACT
AACTTATAAAAATTGCTAACCTCAGCATGTTAAACAATTAAAAGACTGAAAGTCATAGA
TCTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTTGGCTCTGCTCAAATG
CCAGAACTTCTGTAGAAAGTTATGAAACCCAGGTCTGGAACAATTACATTATTCAGATAT
GATAAGTATGCAAGGAGTTGCAGATTCAAAACAAAGAGGCTTCTCATGTCGTTAATGA
AAGCTGCTACAAGTATGGGAGACCTGGATCTAAGTAAAAATAGTATATTGGCTCTGCAAGT
CCTCTGATTTCAGCATCTTCTTCAAATGCCATTACATTGAGATATTGGACTCTCCAA
CAAACCGCTTGATTACTCCATTCAACAGCATTGAAGAGCTCACAAACTGGAAGTTCTGG
ATATAAGCAGTAATAGCCATTATTTCAATCAGAAGGAATTACTCATATGCTAAACTTACC
AAGAACCTAAAGGTCTGCAGAAACTGATGATGAAACGACAATGACATCTTCCCTCACCAG
CAGGACCATGGAGAGTGAGTCTCTAGAACACTCTGGAATTCAAGAGGAATCACTTAGATGTT
TATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAGAATCTGCTAAAATTAGAGGAA
TTAGACATCTCTAAAATTCCCTAAGTTCTGCCCTCTGGAGTTTGATGGTATGCCCTCC
AAATCTAAAGAATCTCTTGGCCAAAATGGGCTCAAATCTTCAGTTGGAAAGAAACTCC
AGTGTCTAAAGAACCTGGAAACTTGGACCTCAGGCCACAACCAACTGACCACTGTCCCTGAG
AGATTATCCAACCTGTTCCAGAACGCTCAAGAACATCTGATTCTTAAGAATAATCAAATCAGGAG
TCTGACGAAGTATTTCTACAAGATGCCCTCCAGTTGCGATATCTGGATCTCAGCTCAAATA
AAATCCAGATGATCCAAAAGACCAGCTTCCAGAAAATGTCCTCAACAATCTGAAGATGTTG
CTTTGCATCATATCGGTTCTGTGCACCTGTGATGCTGTGTTGTCTGGTGGTTAA
CCATACGGAGGTGACTATTCCCTACCTGGCCACAGATGTGACTTGTGTTGGGCCAGGAGCAC
ACAAGGGCAAAGTGTGATCTCCCTGGATCTGTACACCTGTGAGTTAGATCTGACTAACCTG
ATTCTGTTCTCACTTCCATATCTGTATCTCTTCTCATGGTATGACAGCAAGTCA
CCTCTATTCTGGGATGTGTTGATATTACCATCTGTAAAGGCCAAGATAAAAGGGGTATC
AGCGTCTAATATCACCAAGACTGTGCTATGATGCTTTATTGTGATGACACTAAAGACCCA
GCTGTGACCGAGTGGGTTTGGCTGAGCTGGTGGCCAAACTGGAAGACCAAGAGAGAAACA
TTTAATTATGTCAGGAAAGGGACTGGTACCAAGGGCAGCCAGTTCTGGAAAACCTT
CCCAGAGCATACAGCTTAGCAAAAAGACAGTGTGATGACAGACAAGTATGCAAAGACT
GAAAATTAAAGATAGCATTTACTTGTCCCATCAGAGGCTCATGGATGAAAAAGTTGATGT
GATTATCTGATATTCTTGAGAACGCCCTTCAGAAAGTCAAAGTCCAGCTCCGGAAAAA
GGCTCTGTGGGAGTTCTGTCCCTGAGTGGCCAACAAACCCGCAAGCTCACCCATACTCTGG
CAGTGTCTAAAGAACGCCCTGGCACAGACAATCATGTGCCCTATAGTCAGGTGTTCAAGGA
AACGGTCTAGCCCTTTGCAAAACACAACGTGCCTAGTTACCAAGGAGAGGCCCTGGC

215 | 237

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDPKHNHIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQLLEIPQGLPPSLQLLSLEANNIFSIRKENTELANIEILYLGQNCYYR
NPCYVSYSIEKDAFLNLTKVLSDLKDNNVTAVPTVLPSLTELKYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNSPLOIPVNAFDALTELKVRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQLFSSLKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDLGTNFKIANLSMFQFKRLKVIDLSVNKIS
PSGDSSEVGFCASNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRNL
STAFEELHKLEVLDISSLNSHYFQSEGITHMLNFTKNLKVLQKLMMNDNDISSLTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLHHNRFLCTDAWWFVWWVNHTEV
YLATDVTCVPGPAHKGQSVISLDLYTCELDLTNLIFSLISISVSLFLMVMMTASHLYFWDV
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPGPQPVLENLSQSICLSSKTVFVMTDKYAKTENFKIAFYLSHQRLMDEKVDVIILIFE
KPFQSKFLQLRKRLCGSSVLEWPTNPQAHPYFWQCLKNALATDNHVAYSQVFKETV

216/237

FIGURE 210A

GGGTACCATCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACACAACAGAAACATGGAA
 AACATGTTCCCTCAGTCGTCAATGCTGACCTGCATTTCTGCTAATATCTGGTCCTGTGA
 GTTATGCGCCGAAGAAAATTTCAGAAAGCTATCCTTGATGAGAAAAAGCAAATGACT
 CAGTTATTGCAGAGTGCAGCAATCGTCGACTACAGGAAGTCCCCAACGGTGGGCAAATAT
 GTGACAGAACTAGACCTGTCGATAATTACATCACACACATAACGAATGAATCATTCAAGG
 GCTGCAAAATCTCACTAAAATAATCTAAACCACAACCCAAATGTACAGCACCAGAACGGAA
 ATCCCAGTATACAATCAAATGGCTTGAATATCACAGACGGGGCATTCTCAACCTAAAAAC
 CTAAGGGAGTTACTGCTTGAAGACAACCAGTTACCCAAATACCCCTGGTTGCCAGAGTC
 TTTGACAGAACTTAGTCTAATTCAAAACAATATACAAACATAACTAAAGAGGGCATTCAA
 GACTTATAAAACTTGAAAAATCTCTATTGGCCTGGAACTGCTATTAAACAAAGTTGCGAG
 AAAACTAACATAGAAGATGGAGTATTGAAACGCTGACAAATTGGAGTTGCTATCACTATC
 TTTCAATTCTCTTCACACGTGCCACCCAAACTGCCAAGCTCCCTACGCAAACCTTTCTGA
 GCAACACCCAGATCAAATACATTAGTGAAGAAGATTCAAGGGATTGATAAATTAAACATTA
 CTAGATTAAAGCGGGAACTGTCCGAGGTGCTTCAATGCCCATTCATGCGTGCCTGTGA
 TGGTGGTGCTTCAATTAAATAGATCGTTGCTTCAAAACTTGACCCAACCTCGATACC
 TAAACCTCTAGCACTCCCTCAGGAAGATTAATGCTGCTGGTTAAAAATATGCCCTCAT
 CTGAAGGGTGTGGATCTTGAATTCAACTATTAGTGGAGAAATAGTCTCTGGGGCATT
 AACGATGCTGCCCGCTTAGAAATACTTGTCTTAACTTTGTCTCACGGCATTGCAATTAA
 CACAGCATATTAAATATTCCAGAAACTTCTCTAAACTTTGTCTCACGGCATTGCAATTAA
 AGAGGTTATGTGTTCCAGGAACCTCAGAGAAGATGATTCCAGGCCCTGATGCGCTCCAA
 CTTATCGACTATCAACTTGGTATTAATTAAAGCAAATGATTCAAACACTTTCCAAA
 ATTTCCTCAATCTGAAATTATTACTTGTCAAGAAACAGAATATCACCCTGGTAAAGAT
 ACCCGGCAGAGTTATGCAAATAGTTCTCTTCAACGTCAATCCGGAAACGACGCTCAAC
 AGATTGAGTTGACCCACATTGAACCTTATCATTTCAACCGTCTTAAATAAGCCAC
 AATGTGCTGCTTATGGAAAAGCCTTAGATTAAAGCCTAACAGTATTCTTCAATTGGCCA
 AACCAATTGAAAATCTTCTGACATTGCTGTTAAATCTGTCTGCAAATAGCAATGCTCA
 AGTGTAAAGTGGAACTGAATTTCAGCCATTCTCATGTCAAATATTGGATTGACAAACA
 ATAGACTAGACTTGATAATGCTAGTGCTCTTACTGAATTGCTCGACTTGGAGTTCTAGAT
 CTCAGCTATAATTACACTATTTCAGAAATAGCAGGCGTAACACATCATCTAGAATT
 AAATTTCACAAATCTAAAGTTAAACTTGAGCCACAACACATTATAACTTTAACAGATA
 AGTATAACCTGGAAAGCAAGTCCCTGGTAGAATTAGTTTCACTGGCAATGCCCTGACATT
 TTGTGGAATGATGATGACAACAGGTATATCTCCATTTCAAAGGTCTCAAGAATCTGACACG
 TCTGGATTATCCCTTAATAGGCTGAAGCACATCCAAATGAAGCATTCTTAATTGCCAG
 CGAGTCTCACTGAACATATAATGATAATATGTTAAAGTTTTAACTGGACATTACTC
 CAGCAGTTCTCGTCTGAGTTGCTTACTACGTGGAAACAAACTACTCTTTAACTGA
 TAGCCTATCTGACTTACATCTCCCTCGGACACTGCTGCTGAGTCATAACAGGATT
 ACCTACCCCTCTGGCTTCTTCTGAAGTCAGTAGTCTGAAGCACCTCGATTAAAGTCCAAT
 CTGCTAAAAACAATCAACAAATCCGCACTGAAACTAAGACCACCAAAATTATCTATGTT
 GGAACATCACGGAAACCCCTTGAATGACACTGTGACATTGGAGATTCCGAAGATGGATGG
 ATGAACATCTGAATGTCAAATTCCAGACTGGTAGATGTCATTGCTGCCAGCCTGGGAT
 CAAAGAGGGAAAGAGTATTGTGAGTCTGGAGCTAACAACTTGTGTTAGCTAAGGTA
 AAAAGGCTACAGTCTCTTCACTGCTTCTTCACTGCTTCTTCACTGCTACTGCAGT
 GATATTATTCTTCACGTTCTTATCACCACATGGTTATGTTGCTGCCCTGGCTCACC
 ATTGTTTACTGGATGTTGGTTATATATAATGTTGTTAGCTAAGGTA
 AAAAGGCTACAGTCTCTTCCACATCCAAACTTCTATGATGCTTACATTCTTATGACAC
 CACTGTTACTGACTGGGTGATAATGAGCTGCGTACCCACCTGAAGAGAGCCGAGACAAA
 ACGTTCTCTTGTCTAGAGGAGAGGGATTGGACCCGGATTGGCCATCATCGACAAACCTC
 ATGCAGAGCATCAACCAAAGCAAGAAAACAGTATTGTTAAACAAAAAATATGCAAAAG
 CTGGAACTTTAAACAGCTTTACTTGGCTTGCAGAGGCTAATGGATGAGAACATGGATG
 TGATTATATTATCCTGCTGGAGCCAGTGTACAGCATTCTCAGTATTGAGGCTACGGCAG
 CGGATCTGAAGAGCTCCATCCTCCAGTGGCTGACAACCCGAAGGAGAAGGCTTGGTTG
 GCAAACCTGTGAGAAATGTGGTCTGACTGAAAATGATTGACGGTATAACAAATATGTA
 ATTCCATTAAGCAATACTAACTGACGTTAAGTCATGATTGCGCCATAATAAGATGCAAA
 GGAATGACATTCTGTATTAGTTATGCTATGTAACAAATTATCCAAAACCTAGTGG
 TTTAAAACAACACATTGCTGGCCACAGTTTGAGGGTCAGGAGTCCAGGCCAGCATAA

217/237

FIGURE 210B

CTGGGTCTCTGCTCAGGGTGTCTCAGAGGCTGCAATGTAGGTGTTACCAAGAGACATAGGC
ATCACTGGGGTACACTCATGTGGTTTCTGGATTCAATTCCCTGGCTATTGCCA
AAGGCTATACTCATGTAAGCCATGCAGCCTCTCCCACAAGGCAGCTGCTTCATCAGAGCT
AGCAAAAAAGAGAGGTTGCTAGCAAGATGAAGTCACAATCTTGTTAATCGAATCAAAAAAG
TGATATCTCATCACTTGGCCATATTCTATTGTTAGAAGTAAACCACAGGTCCCACCAGCT
CCATGGGAGTGACCACCTCAGTCAGGGAAAACAGCTGAAGACCAAGATGGTGAGCTCTGAT
TGCTTCAGTTGGTCATCAACTATTCCCTTGACTGCTGTCCTGGGATGGCCTGCTATCTG
ATGATAGATTGTGAATATCAGGAGGCAGGGATCACTGTGGACCATCTAGCAGTTGACCTAA
CACATCTCTTCAATATCTAAGAACCTTGCCACTGTGACTAATGGTCCTAATATTAAGC
TGTTGTTATATTATCATATATCTATGGCTACATGGTTATATTATGCTGTGGTTGCCTCG
GTTTATTTACAGTTGCTTTACAAATATTGCTGTAACATTGACTTCTAAGGTTAGATG
CCATTTAAGAACGTGAGATGGATAGCTTTAAAGCATCTTACTTCTTACCAATTAAAAA
GTATGCAGCTAAATTGAAGCTTTGGTCATATTGTTAATTGCCATTGCTGAAATCTTAA
AATGAATGAATAAAATGTTCATTTACAAAAAAAAAAAAAA

218/237

FIGURE 211

MENMFLQSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECASNRRLQEVPQTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVQHQGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSLIQNNIYNITKEGISRLINLKNLYLAWNCFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFFPCVPCDGGASINIDRFAFQNLTQLRYLNLSSTSCLRKINAAWFKNM
PHLKVLDFNYLVGEIVSGAFLTMLPRLEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFQIDFKLFQNFSNLEIIYLSERISPLV
KDTRQSYANSSSFQRHIRKRRSTDFFDPHSNFYHFTRPLIKPQCAAYGKALDLNSIFFI
GPNQFENLPDIACLNLANSNAQVLSGTEFSAIIPHVKYLDLTNNRLDFDNASALTSDLEV
LDLSYNSHYFRIAGVTHLEFIQNFTNLKVNLSHNNIYTLDKYNLESKSLVELVFSGNRL
DILWNDDDNRYSIFKGLKNLTRLDSLNLKHIPNEAFLNLPASLTTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFLTDLSDFSSLRTLLLShNRISHLPSGLSEVSSLKHLDLS
SNLLKTINKSALETKTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFITTMVMLAALAHLYWDVWFIFYNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPGLAIID
NLMQSINQSKKTVFVLKKYAKSWNFKTAFYLALQRLMDENMDVIIIFILLEPVLQHSQYLRL
QRICKSSILOWPDPNPKAEGLFWQTLRVNLTENDSRYNNMYVDSIKQY

219/237

FIGURE 212

CCAGGTCCA ACT GCAC CT CGG TT CT ATCG ATT GA ATT CCCC GGGG AT CCT CT AGAG AT CCCT
CGACCTCGACCCACCGTCCGCCAAGCTGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCCTCAGGCCAGTGGCCTGAGGCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGCCACGCCTGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGAGGCACAGGTGGCCCCACCACCCGGAGGA
GCAGCTCCTGCCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCGCCCTGGAGGCACAGGCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGGTGGCAGTGGCGGCACAGAGCACGCCTACCGGCCGGCGTAGGGTGTGCTGT
CCGGGCTCACGGGACCCCTGTCTCGAGTCGTTGTGCAGCGTGTGTACCAAGCCCTTCCTCA
CCACCTGCGACGGCACCGGCCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGCTGGCCCTGCCAGGCCCTCGCTACCGTGCTGCCCGGCTGGAAGAGGAC
CAGCGGGCTTCCCTGGGCCTGTGGAGCAGCAATATGCCAGCCGCATGCCGGAACGGAGGGA
GCTGTGTCAGCCTGGCCCTGCCGCTGCCCTGCAGGATGGCGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCCAGTGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAAGCTGCAGCTGGTGTGGC
CCCAC TGCA CAGC CTGGC TCGC AGGCA CTGGAGC ATGGC TCCC GGACCCGGCAGC CTCC
TGGTGCACTCCTCCAGCAGCTGGCCGCATCGACTCCCTGAGCGAGCAGATTCTCCTTG
GAGGAGCAGCTGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTG
GACTGAGCCCTCACGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCTTCCCTCTTCCCTCCCC
TTCCTCGGGAGGCTCCCAGACCCCTGGCATGGATGGCTGGATCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCCTGGCTACCCCAACGGCATCCAAAGGCCAGGTGGCCCTCAGCTG
AGGGAAAGGTACGAGCTCCCTGCTGGAGCCTGGACCCATGGCACAGGCCAGGCAGCCGGAG
GCTGGGTGGGCCTCAGTGGGGCTGCTGCCCTGACCCCCAGCACAATAAAATGAAACGTGA
AAAAAAAAAAAAAAAAGGGCGGCCGCACTTAGTTATTGCAGCTTATAATGGTTACAAAT
CGACCTGCAGAAGCTTGGCCGCATGGCCAACTTGTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVL LMWLLV LAVGGTEHAYRPGR RVCAVRAHGD PVSE SFVQRVYQPFLTTCDGH RAC
STYRTIYRTAYRSPGLAPARPRYACCPG WKRTSGLP GACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDT CQSDVDEC SARRGGC PQRCINTAGSYWCQCWE GHLSADGTLCVPKG GPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSI LASQALEHGLPD PGSL LVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

221/237

FIGURE 214

GCCAGGCAGGTGGCCTCAGGAGGTGCCTCAGGCCAGTGGCCTGAGGCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGGCCACCATGCCACGCCCTGGCTCC
AGCAGCATCAGAGCAGCCCCGTGGTGGCAGCAAAGTTAGCTTGCTGGCTGGCCGCTGTGA
GGGCTTCGCGTACGCCCTGCGGTGTCCCAGGGCTGAGGTCTCCTCATTTCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGCCGGGAGGGAACTGGCCCCGAGGGAGAGGAACCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCAGGGAGG
CACAGGTGGCCCCACCACCCGGAGGAGCAGCTCTGCCCTGTCCGGGGATGACTGATT
TCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGCCCTGGAGGCACAGGCATGAGGGC
TCTCAGGAGGTGCTGATGTGGCTCTGGTGTGGCAGTGGCGACAGAGCACGCC
CCGGCCCGGCCGTAGGGTGTGTGCTGCCGGCTCACGGGACCCCTGTCTCCGAGTCGTT
TGCAGCGTGTGTACAGCCCTCACCACCTGCGACGGCACCGGCCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCTACCGCCGAGCCCTGGCTGGCCAGGCCTCGCT
CGCGTGCTGCCCGGCTGGAAGAGGACCAGCGGGCTTCCCTGGGCCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGAACGGAGGGAGCTGTGTCCAGCCTGGCGCTGCCGTGCC
GGATGGCGGGGTGACACTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCCAGTGTGGAGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGGTGGCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGAC
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCTCGCAGGC
ATGGGCTCCGGACCCGGCAGCCTGGTGCACCTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTCTCTGGAGGAGCAGCTGGGTCTGCTCCTGCAAGAAAGA
CTCGTGACTGCCAGCGCTCCAGGCTGGACTGAGCCCTCACGCCGCC
CCCCGCCAACATGCTGGGGTCCAGAACGCCACCTGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTCCTCCTTCCCTCCCTGGAGGCTCCAGAACCCCTGGCTACCC
GGGCTGGATCTCTGTGAATCCACCCCTGGCTACCCACCC
TCCAAGGCCAGGTGGACCCCTCAGCTGAGGGAAAGGTACGAGCTCC
CCATGGCACAGGCCAGGCAGGCCAGGCTGGTGGGCCTCAGTGGGGCTGCTGC
CCCCAGCACAATAAAATGAAACGTG

222/237

FIGURE 215

MRGSQEVL LMWLLV LAVGGTEHAYRPGR RVCAVRAHGD PVSE SFVQRVYQPFLTTCDGH RAC
STYRTIYRTAYR RSPGLAPARPRYACCPG WKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDT CQSDVDECSARRGGCPQRCINTAGSYWCQCWE GHSL SADGTL CVPKG GPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPD PGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

223/237

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCCAGTGGCCTGAGGCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCAGGCCACCATGGCACGCCTGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGC
CTGGAGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGATGTGGCTCTGGTTGGC
AGTGGCGGCACAGAGCACGCCTACCGGCCGGCTAGGGTGTGCTGTCCGGCTCACG
GGGAC CCTGTCTCGAGTCGTTCGTAGCGTGTGACAGCCCTTCCTACCCACCTGCGAC
GGGCACCGGGCTGCAGCACCTACCGAACCATCTATAAGGACCGCCTACCGCCGAGC
GCTGGCCCTGCCAGGCCCTCGCTACCGTGCTGCCCGGCTGGAAGAGGACCAGC
CTGGGGCTGTGGAGCAGCAATATGCCAGGCCATGCCGAACGGAGGGAGCTGTG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGT
ATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCGTAAACACCGCCGGCAGTTACT
GCCAGTGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGG
CCCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGC
GAGAGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGG
GCCTGGCCTCGCAGGCACTGGAGCATGGCTCCGGACCCGGCAGCCTCCTGGTGC
TTCCAGCAGCTGGCCGCATCGACTCCCTGAGCGAGCAGATT
GGGTCCCTGCTCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTGGACTGAG
TCACGCCGCCCCATGCCCTGCCAACATGCTGGGGTCCAGAACGCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCTTCCCTCTTCCCTCCCCCTC
GCTCCCCAGACCCCTGGCATGGATGGGATGGGATCTCTGTGAATCC
CCCACCCCTGGCTACCCAACGGCATCCAAAGGCCAGGTGGCC
GAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAG
CCTCAGTGGGGCTGCTGCCTGACCCCCAGCACAATAAAATGAAACGTG

224/237

FIGURE 217

MRGSQEVL LMWLLVIAVGGTEHAYRPGRVC A VRAHGD PVSE SFVQRVYQPFLTTCDGH RAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDT CQS DVDECSARRGGCPQRCVNTAGSYWCQCWE GHSL SADGTL CVPKG GPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHS LASQALEHGLPD PGSL LVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKD S

225 / 237

FIGURE 218

GGTTGCCACAGCTGGTTAGGGCCCCGACCACTGGGGCCCTTGTCAAGGAGGAGACAGCCTC
CCGGCCCGGGGAGGAAGTCGCTGCCACCTTGGCTGCCACGTGATTCCCTGGGACGGTC
CGTTCTGCCGTAGCTGCCGGCGAGTTGGGCTCCGTGTTCAAGGCCGGCTCCCCCTTC
CTGGCTCCCTCTCCCGCTGGGCCGGTTATCAGGGAGGAGATTGTCTCCAGGGCTAGCAA
TTGGACTTTGATGATGTTGACCCAGCGGCAGGAATAGCAGGCAACGTGATTCAAAGCTG
GGCTCAGCCTCTGTTCTCTCGTAATCGAAAACCATTGGAGCAGGAATTCAA
TCATGTCTGTGATGGTGGTGGAAAGAAGGTGACACGGAAATGGGAGAAACTCCCAGGCAGG
AACACCTTTGCTGTGATGGCCGCGTCATGATGGCCCGCAAAAGGGCATTTCACCTGAC
CCTTTCCATCCTGGGACATGTACACTCTTCTCGCCTTGAGTGCCGCTACCTGGCTG
TTCAGCTGTCTGCCATCCCTGTATGGCTGCCATGCTCTCCTTCTCCATGGCTACA
CTGTTGAGGACCAGCTCAGTGACCCCTGGAGTGATTCCCTGGGCGTACCATGAAAGCAGC
TTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCAGGCGCCAGGGCAGCACCACGC
CTCGTATCAAGAATTTCAGATAAAACACAGATTGTGAAACTGAAACTGTACACATGC
AAGATCTCCGGCTCCCCGGGCCTCCATTGCAAGCATCTGTGACAACACTGTGTGGAGCCTT
CGACCATCACTGCCCTGGTGGGAATTGTGTTGGAAAGAGGAACCTACCGCTACTTCTACC
TCTTCATCCTTCTCTCCCTCACAATCTATGTCCTCGCCTCAACATCGTCTATGTG
GCCCTCAAATCTTGGAAATTGGCTTCTGGAGACATTGAAAGAAACTCTGGAACTGTTCT
AGAAGTCTCATTGCTCTTACACTCTGGTCCGTGTTGGACTGACTGGATTTCATACTT
TCCTCGGGCTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAAGAT
CGCGTCAGAATCCCTACAGCCATGGCAATTGTGAAGAACTGCTGTGAAGTGCTGTGG
CCCCCTGGCCCCCAGTGTGCTGGATCGAAGGGTATTTGCCACTGGAGGAAAGTGGAAAGTC
GACCTCCCAGTACTCAAGAGACCAGTAGCAGCCTTGCACAGAGCCCAGCCCCACAGAA
CACCTGAACCTCAAATGAGATGCCGGAGGACAGCAGCACTCCGAAGAGATGCCACCTCCAGA
GCCCGAGGCCACCACAGGAGGCAGTGAAGCTGAGAAGTAGCCTATCTATGGAAGAGACT
TTTGTGTTGTTAATTAGGGCTATGAGAGATTTCAGGTGAGAAGTTAAACCTGAGACAGAG
AGCAAGTAAGCTGCCCTTTAACTGTTTCTTGGTCTTAGTCACCCAGTTGCACACTG
GCATTTCTGCTGCAAGCTTTTAAATTCTGAACCTAAGGCACTGGCAGAAGATGTCAG
TCACCTCTGATAACTGGAAAAATGGGTCTTGGGCCCTGGCACTGGTCTCCATGGCCTCA
GCCACAGGGTCCCCCTGGACCCCCCTCTTCCCTCAGATCCCAGCCCTCCTGCTTGGGTC
ACTGGTCTCATTGGGCTAAAGTTTGAGACTGGCTCAAATCTCCAGCTGCTGCA
CGTGTGAGTCCAGAGGCAGTCACAGAGACCTGGCCAGGGATCTAACGGTTCTTGG
GGTCTCAGGACTGAAGAGGGAGAGTGGGTAGAAGATTCTCTGGCCACCAAGTGC
AGCATTGCCACAAATCTTTAGGAATGGACAGGTACCTCCACTGTTGTANNNNNNNN
NNNNNNNNNNNNNNNNNTGTTTCTTTGACTCTGCTCCATTAGGAGCAGGAATG
GCAGTAATAAAAGTCTGCACTTGGTCAATTCTTCTCAGAGGAAGCCCGAGTGCTCACT
TAAACACTATCCCTCAGACTCCCTGTGAGGCTGCAGAGGCCCTGAATGCACAAATGGG
AAACCAAGGCACAGAGAGGGCTCCCTCTCCTCTCCCCGATGTACCTCAA
AAAAATGCTAACCAAGTTCCATTAGCCTGGCTGAGTGAGGGAAAGCCAGCAGTGTG
CCCTCTGGGTAACTCACCTAAGGCTCGGCCACCTCTGGCTATGGTAACCACACTGGGG
GCTTCCCTCAAGCCCCGTCTTCCAGCACTCCACCGGCAGAGTCCAGAGGCCACTTCACCC
TGGGGTGGGCTGTGGCCCCAGTCAGCTGCTCAGGACCTGCTCTATTTCAGGGAGAAG
ATTATGATTATATGTGGCTATATTCTAGAGCACCTGTGTTTCTCTTCAAGCCAG
GGTCCTGTCTGGATGACTTATGCCGTGGGGAGTGTAACCGGAACCTTCATCTATTGAA
GGCGATTAAACTGTGTCTAATGCA

226/237

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLILGTCTLFFAFECRYLAV.
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVI~~P~~RALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNYRYFYL
FILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLIKFTLWSVVGLTGHTF
LVALNQTTNEDIKGSGWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPQEAAEAEK

227/237

FIGURE 220

AAAACCCGTATTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTGAAATTAGGTAT
TATAGGGATGGTGGGTTGATTNTTCCTGGAGGCTTTGGCTTGGACTCTCNCTTCT
CCCACAGAGCNCTCGACCACACTGCCCTGGTGGGAATTGTGTTGGAAAGAGGAACTA
CCGCTANTTCTACCTTCACTCCTTNTCTCTCCNCCTCACAACTATGTCTCGCCTTCA
ACATCGT

228/237

FIGURE 221

GTTGTGTCCTTCAGCAAAACAGTGGATTAAATCTCCTTGACAGCTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAACCGAACCTGACAAAAAAGAAGAAAAGAAGA
AAAAAAATCATGAAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGGCAATCTCAC
GGGGCTGGCTGCTCTGTCTCTCCAAGGAGTGCCCCTGCGCAGCGGAGATGCCACCTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGCGCCACCCCTCAGGTGACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGCTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTCAAGTATCTCCAAAATTGTAGAGATTTC
TTCAGATATCTCCATTAATGAAGGAAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCAAAGCGGTTGGCTTGTGAGTGAAGAC
GAATACTTGGAAATTCAAGGCATCACCCGGGAGCAGTCAGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCCGCCGTGGTACGGAGAGTAAAGGTACCGTGAACATCCACCATACA
TTTCAGAACGCAAGGGTACAGGTGCCCCGTGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTCCCTCTCAAAACTCATTTCTCAATGTCTTG
AACATGACTATGGAAACTACACTTGCCTGGCCCTCAAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTGGTCCAGGCGCCGTAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTCGTCTGGCTGCTGCCCTTCTGGTCTTGCACCTGCTTCTCAAATTTGATGTGAGTGCC
ACTTCCCCACCCGGAAAGGCTGCCGCCACCACCAACACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATACAAATGAAATTAGAAGAAACACAGCCTCATGGACAGA
AATTGAGGGAGGGAAACAAAGAATACTTGGGGGAAAGAGTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTAGGTACAATGGAGTTCTTCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCACCTCTTGGTGCAGTGTGGCAA
GGGCTCAGCCTCTGCCACAGAGTCCCCACGTGGAACATTCTGGAGCTGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTCCGGCCAAGCGTGGCGCTCGG
GCACTTGGTAGACTGTGCCACCACGGCGTGTGAAACGTGAAATAAAAGAGCAAAA
AAAAAA

229/237

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVLLSNTQTQYSIEIQNVDVYDEGPYTCVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAvgFVSEDEYL
EIQGITREQSGDYECASNDVAAPVVRVKVTVNYPPISEAKGTGVPVGQKGLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNLGHTNASIML
FGPGAVSEVSNGTSSRAGCVWLLPLLVHLKLKF

230/237

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTCTCTTCCAAGGAGTGCCCGTGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCCTCAGGTGCACTA
TTGACAACCGGGTCAACCGGGTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTCTGAGAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGAAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

231 | 237

FIGURE 224

ATGGCTGGTACGGCGGGGCCGGCAGGGGACCGGGGCCGCGGCCGGAGCAGGCCAGCTG
CCGGGAGCCCTGAATCACCGCCTGGCCCAGTCACCATGAACGTCACGCTGCAGGAGCTGG
GAGCTGGCAGCAACGTGGATTCCAGAAGGGACAAGACAGCTGTTAGGCTCACGCACGCAG
CTGGAGCTGGTCTTAGCAGGTGCCCTCTACTGCTGGCTGCACTGCTCTGGCTGCCCTGT
GCCCTAGGGGTCAGTACACAGAGACCCATCCCACAGCACCTGCCCTACAGAGGCCTGCA
TTCGAGTGGCTGGAAAAATCCTGGAGTCCCTGGACCGAGGGTGAGGCCCTGTGAGGACTTT
TACCAGTTCTCCTGTGGGGCTGGATTCGAGGAACCCCTGCCGATGGCGTTCTCGCTG
GAACACCTCAACAGCCTCTGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACA
CCACCTCAACTCCAGCAGTGAAGCTGAGCAGAAGACACAGCGCTTACCTATCTGCTA
CAGGTGGAGCGCATTGAGGAGCTGGGAGGCCAGCCACTGAGAGACCTCATTGAGAAGATTGG
TGGTTGGAACATTACGGGCCCTGGGACCAGGACAACTTATGGAGGTGTTGAAGGCAGTAG
CAGGGACCTACAGGGCCACCCATTCTCACCGTCTACATCAGTGCCACTCTAACAGATTCC
AACAGCAATGTTATCAGGTGGACCAGTCGGCTCTTCTGCCCTCGGGATTACTACTT
AAACAGAACTGCCAATGAGAAAGTGCTACTGCCCTATCTGGATTACATGGAGGAACGGGG
TGCTGCTGGGTGGCGGGCCACCTCCACGAGGGAGCAGATGCAGCAGGTGCTGGAGTTGGAG
ATACAGCTGGCAACATCACAGTCCCCAGGACCAGCGCGACGAGGAGAAAGATCTACCA
CAAGATGAGCATTCGGAGCTGCAGGCTCTGGGCCCTCCATGGACTGGCTTGAGTTCTGT
CTTCTTGCTGTCACCATTGGAGTTGAGTGACTCTGAGCCTGTGGTGTATGGATGGAT
TATTGAGCAGGAGTGTCAAGAGCTCATCAACCGCAGGAACAGCATCCTGAACAAATTACCT
GATCTGGAACCTGGTCAAAAGACAACCTCAAGCCTGGACCGACGCTTGAGTCTGCACAAG
AGAAGCTGCTGGAGACCCCTATGGCACTAAGAAGTCCTGTGCGAGGTGGCAGACCTGC
ATCTCCAACACGGATGACGCCCTGGCTTGCTTGGGACTCTTGTGAAGGCCACGTT
TGACCGGCAAAGCAAAGAAATTGAGAGGGATGATCAGCGAAATCCGGACCGATTGAGG
AGGCCCTGGGACAGCTGGTTGGATGGATGAGAAGACCCGCCAGGAGCCAAGGAGAAAGCA
GATGCCATCTATGATATGATTGGTTCCAGACTTTATCCTGGAGCCAAAGAGCTGGATGA
TGTTATGACGGGTACGAAATTCTGAAGATTCTTCTCCAAACATGTTGAATTGTACA
ACTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAGCCTCCCAGCCGAGACCAGTGGAGC
ATGACCCCCCAGACAGTGAATGCCACTACCTTCAACTAAGAATGAGATCGTCTCCCCGC
TGGCATCTGCAGGCCCTTCTATGCCGCAACCACCCAAAGGCCCTGAACCTCGGTGGCA
TCGGTGGTCACTGGCCATGAGTTGACGCATGCCCTTGATGACCAAGGGCGAGTATGAC
AAAGAAGGGAACCTGCGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTCCGGAACACAC
GGCCTGCATGGAGGAACAGTACAATACCAAGGTCAATGGGAGAGGCTAACGGCCGCC
AGACGCTGGGGAGAACATTACTGACAACAGGGGGCTGAAGGCTGCCCTACAATGCTACAAA
GCATGGCTGAGAAAGCATGGGAGGGAGCAGCAACTGCCAGCCGTGGGCTACCAACCACCA
GCTCTCTCGTGGGATTGCCAGGTGTTGCTCGGTCCGCACACCAGAGAGCTCTCACG
AGGGGCTGGTGAACGCCACAGCCCTGCCGCTCCCGTGTGGCAGTCTCTCAAC
TCCCGTGAATTCTCGGCCACTTGGCTGCCCTGTCGGCTCCCCATGAACCCAGGGCAGCT
GTGTGAGGTGTTTAGACCTGGATCAGGGAGAAATGGCAGCTGTCACCAGACCTGGGCA
GCTCTCTGACAAAGCTGTTGCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGCTGGGT
CTAGTCCCTCCCCCAGGGTACATGAGTACAGACCCCTCTCAATACCAATTGTGCCT
CTGCTTGGGGTGCCCTGCCCTGTCAGCAGAGCCCCACCATTCACTGTGACATCTTCCGT
GTCACCTGCCTGGAAGAGGTCTGGTGGGAGGCCAGTCCCATAGGAAGGAGTCTGCC

232/237

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLALLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPPLPDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNIITGPWDQDN
FMEVLKAVAGTYRATPFFTVDYISADSKSSNSNVIQVDQSGFLFLPSRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLNSRDFLRHFGCPVGSPMNPGQLCEVW

233 | 237

FIGURE 226A

GCCCGGCCCTCCGCCCTCCGCACTCCGCCTCCCTCCGCCGCTCCGCCCTCCTC
 CCTCCCTCCTCCCCAGCTGTCCCCTCGCTCATGCCGAGCCTCCGCCGCCGGCCCC
 CTGCTGCTCCCTCGGGCTGCTGCTCGCTCGGCCCTCCGCCGGCCGCCGGCCCC
 CCCCGTGTGCCATCCGTTCTGAGAAGGAGGCCGTGCCCTCGGGAGCAGGTAGGT
 GGGCGCCCAGGGGAGGCGCGGGCGGGAGTCGGGCTCGGGCGAGTCAGGCCAGCCCC
 GGGCGCGGGGGCAGGTGGCTCGGCCGGCGGGGGAGGGTGGCGGGGGAGAAG
 GGCGCGGTGCCTGGGACCCGGACCGCAGGCCACACGGCGCAGCTG
 GGCAGCGGCCTCCAGCCAAGGCCGTCCCCCAGGTGACACCTTCGGCGGGAGGTCTATGCC
 TTGGACGAGACGTGGCACCCGACCTAGGGAGCCATTGGGTGATGCGCTGCGTGTG
 CGCCTGCGAGGCCAGTGGGTGCGTACCAAGGGCCCTGGCAGGGTCAGCTGCAAGAAC
 TCAAACCAAGAGTGCCTAACCCGGCTGTGGCAGCCGCAGCTGCCGGACACTGCTGC
 CAGACCTGCCCCCAGGACTTCGTGGCGCTGTCAGACAGGGCCAGGTGCGAGGCCGGTGGCACG
 AGCCCAGTCTCGCTGCGCTTAGCCTCCGCTCTATCTCTACAGGCCGTGGACC
 GCCCTACCAGGATCCGCTCTCAGACTCCAATGGCAGTGTCTGTTGAGCACCCCTGCAAGCC
 CCCACCAAGATGGCCTGGCTGTGGGTGTGGCAGGCCAGTCGGTTGTCTGCGGCT
 CCTTAGGGCAGAACAGCTGCATGTGGCACTTGTGACACTCACTCACCTTCAGGGAGGTCT
 GGGGGCCTCTCATCCGGCACCGGGCCCTGTCCCCAGAGACCTTCAGTGCCTACTGACTCTA
 GAAGGCCCCACCAGCAGGGCGTAGGGGGCATCACCTGCTCACTCTCAGTGCACACAGAGGA
 CTCCTTGCAATTGGCTGCTCTCCGAGGCCTTGCAAGGACTAACCAAGGTTCCCTGAGGC
 TCCAGATTCTACACCAGGGGAGCTACTGCGAGAACCTCAGGCCATGTCTCAGGCCAGGAA
 CCAGGCTTGCTGAGGTGCTGCCAACCTGACAGTCCAGGAGATGGACTGGCTGGTGTGG
 GGAGCTGCAGATGGCCTGGAGTGGCAGGCAGGGCTGCGCATCAGTGGACACATTG
 CTGCCAGGAAGAGCTGCGACGTCTGCAAAGTGTCTTGTGGGCTAATGCCCTGATCCA
 GTCCAACGGGTGCTGCCGGCTCAGCCAGCCTCACTCTGCTAGGAAATGGCNCCTGATCCT
 CCAGGTGCAATTGGTAGGGACAACCAGTGAGGTGGGAGGCTGACACTGGAAACCAAGCCTC
 AGCGGAGGGATCAGCCCAGTGTCTGTGCCACATGGCTGGCTATCCTCCCTGCCAGG
 CGTGGGTATCTGCCCTGGCTGGGCTGCCAGGGGCTCATATGCTGCTGAGAACATGAGCT
 CTTCTGAACGTGGCACCAAGGACTTCCAGCAGGGAGCTTCGGGCAACGTGGCTGCC
 CTGCCCTACTGTGGGCACTAGGCCCGCCCTGCCGTGCCCTAGCAGGAGCCCTGGTGT
 CCCCCTGTGAAGAGCCAAGCAGCAGGGCACGCCCTGGCTTCCTGGATAACCACTGTCACCT
 GCACTATGAAGTGTGCTGGCTGGCTGGCTGAGAACAAAGGACTGTCACTGCCACC
 TCCTTGGCCTCTGGAACGCCAGGGCTGGCGCTGCTGAAGGGATTCTATGGCTCAGAG
 GCCCAGGGTGTGGTAAGGACCTGGAGCCGAAGTGTGCGGCACCTGGAAAAGGCATGGC
 TTCCCTGATGATCACCAAGGTAGCCCCAGAGGGAGCTCGAGGGCAGCCTCTCCTCCC
 AGGTGACATAGCCAACCAATGTGAGGTTGGCGACTGCGCTGGAGGGCCGGCCGAG
 GGGGTGCGGGCGCTGGGGCTCCGGATACAGCCTCTGCTGCCGCCCTGTGGCTGGTCT
 CCCGGCCCTAGGCCCGCCAAACCTGGTGGCTCTGGCGGCCGGAGACCCCAACACATGCT
 TCTTCGAGGGGAGCAGCGCCCCACGGGGCTCGCTGGCGCCCAACTACGACCCGCTG
 TCACTCTGCACCTGCCAGAGACGAACGGTATCTGACCCGGTGGTGTGCCACCGCCAG
 CTGCCACACCCGGTGCAGGCTCCGACCAGTGCTGCCCTGTTGCCCTGGCTGCTATTG
 ATGGTGACCGGGAGCTGGCGGGAGCAGCGGGTACCGGGTGGCACCCGTTGTGCC
 TTAATTAAAGTGTGCTGCAACCTGCAAGCAGGGGGCAGTGGAGAGGTGCACTGTGAGAA
 GGTGAGTGTCCCCGGCTGGCTGTGCCAGCCTGTGCGTGTCAACCCACCGACTGCTGCA
 AACAGTGTCCAGGTGAGGCCACCCAGCTGGGGACCCATGCAGGCTGATGGCCCCGG
 GGCTGCCGTTTGCTGGCAGTGGTTCCAGAGAGTCAGAGCTGGCACCCCTCAGTGCC
 GTTGGAGAGATGAGCTGTATCACCTGCAAGATGTGGGTAAGTGGGAGCAGAGGCTGT
 GAGGTGGTACTGGGAGGCTGGCTGGAGTAGGGAGACCTTCCCAGGGAGGTCCCTGAAGAA
 GCTGAAGGTACTGTGCTCCAGTGCTCTGGGGACACTCAGTGTCTGCTCTGT
 AGGCAGGGTGCCACTGTGAGCGGGAGTGAAGTGTCACTGCCACTGCTCTGGCTGG
 AAGGAGAGTCGATGCTGTTCCCGCTGCACGGCCCACCGCGCGTAAGTGGAGGAGTC
 GTCAGCAGCTGTGAGTGGAGGGCTCACCTGCCGTGGACTCCTGATCAGGAAAGGGAGC
 TCACTGTGCAAGAACAGTGCAAGCCTGCCACAAAGTGCCTACAAATCCACCCCTCACAGC
 AACCTGGTGAATTGTTATTGACCTTACAATGAGATTCTGAAGCTCAGAGA
 AATTAAGCAACGAGATGAAGGTACCCAGCTGTGCAGTGCACCTGTTAGAAAATACTGGC

234/237

FIGURE 226B

CTTTCTGGGACCAAGGCAGGGATGCTTGCCCTGCCCTCATGCCTCTGTGCCTCTCCAC
TCCCCTCCCCCTCCTCCAACATTCCCTCCCTCTGCTCTCCAGCAGCCCCAGAGACCAGAACT
GATCCAGAGCTGGAGAAAGAAGCCGAAGGCTCTAGGGAGCAGCCAGAGGGCCAAGTGACCA
AGAGGATGGGGCCTGAGCTGGGAAGGGTGGCATCGAGGACCTCTGCATTCTCCTGTGG
GAAGCCAGTGCCCTTGCTCCTCTGCTCTACTCCCACCCCCACTACCTCTGGGAAC
CACAGCTCCACAAGGGGAGAGGCAGCTGGGCCAGACCGAGGTACAGCCACTCCAAGTCCT
GCCCTGCCACCCCTGGCCTCTGCTCTGGAAGCCCCACCCCTTCTTCTGTACATAATGTCA
CTGGCTGTTGGATTTTAATTATCTTCACTCAGCACCAAGGGCCCCGGACACTCCACTC
CTGCTGCCCTGAGCTGAGCAGAGTCATTATTGGAGAGTTGTATTATTAAAACATTCT
TTTCAGTCTTGGCATGAGGTGGCTCTTGTGCCAGGAACCTGAGTGGGGCTGGTGG
AGAAGGGGNGAGAGTAGGAGGTGAGAGAGAGGAGCTCTGACACTTGGGGAGCTGAAAGAGA
CCTGGAGAGGCAGAGGATAGCGTGGCNNTGGCTGGCATNCCTGGGTCCGAGAGGGCTG
GGGATGTTCTTGAGATGGCTAGAGACTCAAGAATTAGGAAAGTAGAACAGCAGGATTGAG
CTCAAGTTAGTTCCACATCGCTGGCCTGTTGCTGACTCATGTTGAAGTTGCTCCAG
AGAGAGAATCAAAGGTGTCACCAGCCCCCTCTCCCTCCCTCCCTTCCCTTCC
CCTCCCCCTCCCCCTCCCCCTCCCCCTCC

235/237

FIGURE 227

GGCGGAGCGGGGGTGTGCGCGGCCGTATGGCTGGTGACGGCGGGCCGGCAGGGGA
CCGGGGCCGCGGCCGGGAGCGGGCCAGCTGCCGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTCGCCTGCAGGAGCTGGAGCTGGCAGCAACGTGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTAGCAGGTGCCTCTAC
TGCTGGCTGCACTGTTCTGGCTGCCTTGTGGCCCTAGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCGTGAGGACTTTACCAAGTTCTCCTGTGGGGCTGGATTGGGA
GGAACCCCCCTGCCGATGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTGAAAACACCACCTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTGCCTACAGGTGGAGCGCATTGAGGAGCTGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTGGAACATTACGGGCCCTGGACCAAG
GACAACTTATGGAGGTGTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTCAC
CGTCTACATCAGTGCCACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTTTCTGCCCTCTGGGATTACTACTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGCTGATCCCTGTTGACTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCT
CCTTCTTCTTCTTCTCCCTCCCTCCCTTTCTCCCTTTCTCCCTTCCCTTCC
TCTTATTCTCTAGTAGGTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGAAATTG
GAGATATAAGTTCCGAGCCATTGCCACAGGAAGCGTCAGTGTGATGGGTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAGAGGGCCTGAGGATTAGGAGAGACTTATGGAGCC
AGCAAAGTCTCCTGAAGAGATTGCATTGAGCCAGGTCTGTAG

236/237

FIGURE 228

ATGCCTACTACCTTCCA ACTAAGAATGAGATCGTCTCCCCGTGGCATCCTGCAGGCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACCTCGGTGGCATCGGTGTGGCATGGCCA
TGAGTTGACGCATGCCTTGATGACCAAGGGCGCGAGTATGACAAAGAACGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTCCGAAACCACACGGCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGAGAACAT
TGCTGACAACGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGCTACCAACCACAGCTCTCAGAGGGCTGGTGACCGACCC
GCCAGGGTGTGGTGCTCGTCCGCACACCAGAGAGCTCTCACGAGGGCTGGTGACCGACCC
CCACAGCCCTGCCGCTTCCGCGTGTGGCACTCTCCAACCTCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTGGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGCCAGCTGTCAACCAGACCTGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTGGTTGGGAGGAAGCAAATGCAAGCTGGCTGGGTCTAGTCCCTCCCCCACA
GGTGACATGAGTACAGACCCCTCCTCAATCACCACATTGTGCCCTGCTTGGGGTGCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTCACCTGCCCTGGAAGAG
GTCTGGGTGGGAGGCCAGTCCCATAGGAAGGAGTCTGCCCTTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGCCTGCCGTGCCTGCCACTGTGACCCACAGGCCCTGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCACACTAGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTCACGGCTACCCCCACCTCACCCGTGCTCCTGTGCCACTGCTCCCAGTG
CTGCTGCTGACCTTCACTGACAGCTCTAGTGGAAAGCCAAGGGCTCTGAAAGCCTCCTGC
TGCCCAC TGTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCTGT
GTCTTAGGGCACAAGCCTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTATTTTACAGAAAAGAGGGTGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

237 / 237